Investigation of the sequences and structural elements required for HIV-2 infectivity and genome dimerisation

Thesis

Investigation of the Sequences and Structural Elements Required for HIV-2 Infectivity and Genome Dimerisation

Anne L'Hernault

The Open University

Thesis submitted for the degree of Doctor of Philosophy at the Open University

Registered discipline: Life Sciences

Submission date: 11.06.09

Submission date: 22 June 2006
Date of award: 16 Sept. 2009

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Abstract

A unique feature of retroviruses is that two copies of the genomic RNA are packaged in each particle. The selective encapsidation of viral genomes is ensured by the binding of the Nucleocapsid to a specific motif on the RNA genome, the packaging signal (Psi). The Psi regions of many retroviruses overlap with sequences that promote the dimerisation of the genome, the dimerisation initiation site (DIS), and it has been suggested that the two mechanisms are closely linked.

The aim of the research presented herein was to identify the sequences and structural elements required for the dimerisation of HIV-2 genomic RNA and to investigate the relationship between HIV-2 genome dimerisation and encapsidation, infectivity and particle morphogenesis.

Mutations of two palindromic sequences, introduced in an infectious molecular clone of the HIV-2_{ROD} isolate, revealed that a palindrome within HIV-2 Psi was important for genome dimerisation. In contrast with previous studies, the palindrome termed DIS is not required for genome dimerisation and viral replication.

Viruses bearing mutations within the Psi region failed to dimerise and to replicate in T-cells, a defect that could not be rescued by targeting more genomes to the cells. Psi-deleted viruses also displayed a defect in particle morphogenesis. A reduced packaging efficiency, combined with the presence of RNA monomers or unstable dimers in these virions, resulted in the production of fewer mature particles. However an increase in the number of particles containing two cores was observed.

Further characterisation of the sequences and structural elements required for RNA dimerisation, packaging and viral replication showed that the formation of stem B is not critical for viral replication. However, a GGAG purine-rich motif at position 392-395 of the HIV-2_{ROD} genome is absolutely essential for genome dimerisation and viral infectivity, and a correlation was observed between dimer formation and viral replication.

Declaration

This dissertation is the result of my own work and includes nothing that is the outcome of work done in collaboration, with the exception of the electron microscopy experiment described in Chapter 5, which was performed at the MRC-LMB Cambridge in collaboration with Dr. R.A. Crowther.

Acknowledgments

First, I'd like to thank my supervisor, Andrew Lever, for all his help, time and support, as well as his constant faith in my abilities. Thank you for letting me decide on how to carry on the project, allowing me to make my own mistakes and trying to keep me focus. Of course, science is nothing without money, so thank you for providing the funding and allowing me to go to more conferences than I could have dreamt of.

I'd also like to thank my second supervisor, Jane Greatorex, for teaching me the secret ways of RNA and all I know about dimerisation, being patient, rescuing sequencing gels, showing me everything she knew and letting me find out the rest by myself, giving me the independence I sought and always trusting me, for all the useful discussions and for reading numerous thesis drafts.

There are many people to thank in the Lever/Sinclair lab and I'll try not to forget anyone. Thanks to Harriet for going through this with me, for conference trips and coping with my non-veganess. Thanks to the girls in the office, Emma A., Emma P., Jing, Julia and Truus for scientific discussions, gossips and lots of fun. Big thanks to Joan for making the lab a better place and for the fantasy football. Thanks to Matt and Ian for making the lab a fun place, for Friday nights in the Flee, the fantasy football, teaching me what being English means and reminding me every day that I'm French. Thanks to Ed for insights into British politics and teaching me FACS. Thanks to Neil and Mark Wills for their help with the FACS, to John Sinclair for useful scientific discussions and the cryptic crosswords, to Gavin for encouragements and cheering me up and to Claire for organising the lab. Thanks to Steve Griffin for starting this whole project, help with the infamous RPAs and a few figures.

I'd also like to acknowledge Dr. Tony Crowther for his help with the electron microscope and Dr. Ignacio Romero for useful discussions.

A second thanks to Andrew and Jane, without whom none of this would have happened and to Harriet, Emma A., Julia and Truus for their help at various stage of my PhD.

Last but not least, thanks to Fred for being there when I needed him, to my brother for his encouragements and to my mother for believing in me no matter what, supporting me throughout and being an inspiration.

Table of Contents

List of Figures

List of Tables

Abbreviations

aa	Amino acid
ADP	Adenosine diphosphate
AIDS	Acquired immunodeficiency syndrome
AIP1/Alix	ALG-2-interacting protein 1
AGM	African green monkey (*Chlorocebus pygerythrus)*
Amp	Ampicillin
APOBEC3G	Apolipoprotein B mRNA-editing enzyme catalytic subunit 3G
AP	Adaptor protein
Arf	ADP ribosylation factor
ART	Antiretroviral therapy
ARV	AIDS-associated retrovirus
ATCC	American Type Culture Collection
ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
BMH	Branched structure with multiple hairpin
bp	base-pair
BSA	Bovine serum albumin
Bst-2	B cell stromal factor 2
°C	Degree(s) centigrade
C	Carboxy
CA	Capsid
CCR	CC-chemokine receptor
CD	Cluster of determinants
CDK	Cyclin-dependent kinase
CFAR	Centre for AIDS Reagents
CHMP4	Chromatin modifying protein 4
cpm	Count(s) per minute
cPPT	central polypurine tract
CPZ	Chimpanzee (*Pan troglodytes troglodytes*)
Crm1	Chromosome region maintenance 1

Cryo-EM	Cryo-electron microscopy
CTD	Carboxy-terminal domain
CTL	Cytotoxic T lymphocyte
Cul4A	Cullin 4A
CXCR4	CXC-chemokine receptor 4
CypA	Cyclophilin A
DC	Dendritic cell
DCAF1	DDB1-Cul4A-associated factor
DDB1	Damaged-DNA-specific binding protein 1
DEAE	Diethylaminoethyl
DEPC	Diethylpyrocarbonate
DIS	Dimer initiation site
DLS	Dimer linkage site
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
ECACC	European Collection of cell cultures
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EIAV	Equine infectious anaemia virus
eIF5B	Eukaryotic initiation factor 5B
EM	Electron microscopy
EMSA	Electrophoretic mobility shift assay
Env	Envelope
ESCRT	Endosomal sorting complexes required for transport
FACS	Fluorescence activated cell sorting
FBS	Fetal bovine serum
FDC	Follicular dendritic cell
FIV	Feline immunodeficiency virus
FRET	Fluorescence resonance energy transfer
g	Gravity
Gag	Group-specific antigen
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase

GFP	Green fluorescent protein
GGA	Golgi-localised γ-ear containing Arf-binding
GOR	Gorilla *(gorilla gorilla)*
gp	Glycoprotein
GST	Glutathione S-transferase
h	Hour(s)
HAART	Highly active antiretroviral therapy
HaSV	Harvey sarcoma virus
hCMV	Human Cytomegalovirus
HIV	Human immunodeficiency virus
HIV-1	Human immunodeficiency virus type 1
HIV-2	Human immunodeficiency virus type 2
hnRNP	Heterogeneous nuclear ribonucleoprotein
HPLC	High performance liquid chromatography
HRP	Horseradish peroxidase
HTLV	Human T-cell leukaemia virus
IAA	Isoamyl alcohol
IFN	Interferon
IN	Integrase
IPTG	Isopropyl-ß-D-thiogalactopyranoside
IRES	Internal ribosome entry site
kb	Kilobase(s)
kDa	Kilodalton
KIF4	Kinesin family member 4
LAV	Lymphadenopathy associated virus
LB medium	Luria Bertani medium
LDI	long-distance base-pairing interaction
LTNP	Long-term non progressor
LTR	Long terminal repeat
M	Molar
MA	Matrix
MAC	Macaque (*Macaca*)
MHC	Major histocompatibility complex
MHR	Major homology region

min	Minute(s)
MLV	Murine leukaemia virus
MNE	Pig-tailed macaque (*Macaca nemestrina*)
MVB	Multivesicular body
MW	Molecular weight
N	Amino
NC	Nucleocapsid
NEB	New England Biolabs
Nef	Negative factor
NES	Nuclear export signal
NF-κB	Nuclear factor kappa B
NIBSC	National Institute for Biological Standards and Controls
NK	Natural killer
NLS	Nuclear localisation signal
NMR	Nuclear magnetic resonance
nt	Nucleotide(s)
NTD	Amino-terminal domain
OD	Optical density
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBMC	Peripheral blood mononuclear cell
PbS	Phosphate buffered saline
PBS	Primer binding site
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
$PI(4,5)P_2$	Phosphatidylinositol-(4,5)-biphosphate
PIC	Pre-integration complex
Pol	Polymerase
Pol II	RNA polymerase II
PolyA	Polyadenylation
PPT	Polypurine tract
PR	Protease
Psi	Packaging signal
P-TEFb	Positive transcription elongation factor b

R	Repeat region
RER	Rough endoplasmic reticulum
Rev	Regulator of expression of virion proteins
RH	RNase H
RhMac	Rhesus macaque (*Macaca mulatta*)
RNA	Ribonucleic acid
RNase	Ribonuclease
RPA	Ribonuclease (or RNase) protection assay
rpm	Rotation(s) per minute
RPMI	Roswell Park Memorial Institute
RRE	Rev-responsive element
RSV	Rous sarcoma virus
rt	Room temperature
RT	Reverse transcriptase
RT-PCR	Reverse transcription-polymerase chain reaction
sd	Standard deviation
SD	Splice donor
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
sec	Second(s)
SELEX	Selected evolution of ligands by exponential enrichment
SHAPE	Selective 2′-hydroxyl acylation analysed by primer extension
SIV	Simian immunodeficiency virus
SL	Stem-loop
SM	Sooty mangabeys (*Cercocebus atys*)
SRP	Signal recognition particle
SSC	Standard sodium citrate
STM	Stump-tailed macaque (*Macaca arctoides*)
SU	Surface
SV40	Simian vacuolating virus 40 or Simian virus 40
TAF	TBP-associated factor
TAR	*Trans*-activation responsive
Tat	*Trans*-activator of transcription
TBE	Tris-borate-EDTA

TBP	TATA-binding protein
TCR	T-cell receptor
TEMED	*N,N,N',N'*-tetramethylethylenediamine
T_H	Hybridisation temperature
TIP47	Tail-interacting protein of 47 kDa
T_M	Melting temperature
TM	Transmembrane
TRIM	Tripartite motif protein
TRIS	Tris(hydroxymethyl)aminomethane
tRNA	Transfer ribonucleic acid
Tsg101	Tumor susceptibility gene 101
U3	Unique 3′
U5	Unique 5′
Ub	Ubiquitin
UNAIDS	Joint United Nation programme on HIV/AIDS
UTR	Untranslated region
UV	Ultraviolet
V	Volume
Vif	Virion infectivity factor
VLP	Virus-like particle
Vpr	Viral protein r
Vpu	Viral protein u
Vpx	Viral protein x
WT	Wild type

1 Introduction

1.1 The origins of HIV and AIDS

1.1.1 The discovery of HIV

- ***The identification of AIDS and HIV***

Acquired immunodeficiency syndrome or AIDS was first reported in the early 1980's, when a growing number of previously healthy homosexual patients presented with pneumocystis carinii pneumonia, mucosal candidiasis, Kaposi's sarcoma and multiple viral infections, including human Cytomegalovirus (hCMV) (Friedman-Kien et al., 1982; Gottlieb et al., 1981; Mildvan et al., 1982; Stahl et al., 1982). All patients were found to be lymphopenic and to have an acquired immunodeficiency syndrome.

The etiological agent of AIDS, lymphadenopathy associated virus (LAV), was isolated in 1983 from lymphocytes of a patient presenting with multiple lymphadenopathy (Barre-Sinoussi et al., 1983). LAV was described as a typical type C RNA tumour virus and antibodies from the serum of this patient reacted with proteins from human T-cell leukaemia virus type 1 (HTLV-1). However, specific antibodies against HTLV-1 did not precipitate the proteins of the new isolate.

In the same year, another study reported the isolation of HTLV-1 from patients with AIDS (Gallo et al., 1983). However, two new viruses, named HTLV-III and AIDS-associated retrovirus (ARV), were subsequently isolated from patients with AIDS (Gallo et al., 1984; Gallo et al., 1983; Levy et al., 1984; Popovic et al., 1984; Sarngadharan et al., 1984) and it was later agreed that LAV, ARV and HTLV-III were

the same virus isolate and represented the causative agent of AIDS. LAV/HTLV-III was later renamed human immunodeficiency virus type 1 (HIV-1) as it became apparent that this virus was not related to C-type oncoretroviruses like HTLV (Coffin et al., 1986).

A few years later, another retrovirus was isolated in Western Africa from patients with AIDS (Clavel et al., 1986a). This new retrovirus presented similarities with HIV-1 both morphologically and biologically as its tropism and cytopathic effect on $CD4^+$ T-cells were similar to that of HIV-1 (Brun-Vezinet et al., 1987; Clavel et al., 1986a; Clavel et al., 1986b). Yet the sequence homology at the protein level appeared very limited, with around 50 and 55 % homology for the Gag (group-specific antigen) and Pol (Polymerase) proteins respectively and only 35 % homology for the envelope (Env) glycoprotein (Franchini et al., 1987; Guyader et al., 1987). Serum from a macaque infected with simian T-cell leukaemia virus-type III ($STLV\text{-}III_{MAC}$; also known as simian immunodeficiency virus macaque or SIV_{MAC}; Daniel et al., 1985; Kanki et al., 1985c) could recognise the Envelope glycoprotein from the new retrovirus while serum from an HIV-1 infected patient could not (Clavel et al., 1986a). The serological cross-reactivity was in fact restricted to the major core protein Gag and it was suggested that this new retrovirus, later named HIV-2, was only distantly related to HIV-1. Amino acid sequence comparison confirmed that HIV-2 was more closely related to an SIV isolate from African green monkeys (SIV_{AGM}) than it was to HIV-1 (Franchini et al., 1987). Another human T-lymphotropic virus isolated from apparently healthy Senegalese women and named HTLV-IV showed high similarities and serological cross-reactivity with SIV_{AGM} (Kanki et al., 1986). HTLV-IV was confirmed to belong to the HIV-2 lineage following nucleotide sequence comparison (Hahn et al., 1987).

- ***Nomenclature***

The *Retroviridae* family comprises seven classes of viruses (Fauquet et al., 2005). HIV-1 and -2 belong to the *Lentivirinae* class. Other *Lentivirinae*, named as such due to the slow progression of the disease they cause, include simian, feline and bovine immunodeficiency viruses (respectively SIV, FIV and BIV), Visna Maedi virus (VISNA), caprine arthritis encephalitis virus (CAEV) and equine infectious anaemia virus (EIAV).

The other classes of exogenous *Retroviridae* are the *Alpharetrovirinae* (type species: avian leukaemia virus, ALV and Rous sarcoma virus, RSV), *Betaretrovirinae* (mouse mammary tumour virus, MMTV and Mason-Pfizer monkey virus, MPMV), *Gammaretrovirinae* (murine and feline leukaemia virus, MLV and FLV), *Deltaretrovirinae* (HTLV and bovine leukaemia virus, BLV), *Epsilonretrovirinae* (walleye dermal sarcoma virus, WDSV) and the *Spumavirinae* (human and simian Foamy virus, HFV and SFV).

1.1.2 The emergence of HIV: how, where and when?

Despite allegations that an oral OVP-polio vaccine contaminated by SIV was responsible for the emergence of HIV and the AIDS pandemic (Hooper, 1999), it has now been agreed among the scientific community that it was not the origin of HIV (Berry et al., 2001; Berry et al., 2005). Instead, HIV-1 and -2 are thought to have arisen following several zoonotic transmission events from SIV-infected non-human primates to humans in Africa (Gao et al., 1999; Gao et al., 1992; Hahn et al., 2000; Hirsch et al., 1989; Huet et al., 1990; Keele et al., 2006; Sharp et al., 1995). Five lines of evidence have been used to support the hypothesis of a zoonotic origin of HIV-1 and -2:

similarities in genome organisation, phylogenetic relatedness, prevalence in the natural host, geographic coincidence and plausible route of transmission (Sharp et al., 1995).

- ***Evolution of HIV-2 from the SIV_{SM} lineage***

HIV-2, which was discovered in Western Africa and is only endemic there (Clavel et al., 1986a; Kanki et al., 1986), has been found to be closely related to SIV_{MAC} and SIV_{AGM} due to sequence homology and serological cross-reactivity (Daniel et al., 1985; Franchini et al., 1987; Hahn et al., 1987; Kanki et al., 1985a; Kanki et al., 1985b; Kanki et al., 1985c).

Isolation and characterisation of an isolate of SIV from sooty mangabeys (*Cercocebus atys*), SIV_{SM}, confirmed that HIV-2 was more closely related to the SIV_{SM} lineage than to HIV-1 (Hirsch et al., 1989; Marx et al., 1991; Santiago et al., 2005; Sharp et al., 1995; see Figure 1.1). Sequence comparisons between SIV_{SM}, SIV_{MAC} and HIV-2 suggested that SIV_{SM} had infected humans in Western Africa and captive Rhesus macaques (RhMac; *Macaca mulatta*) and had evolved as HIV-2 and SIV_{MAC}, respectively (Gao et al., 1992; Hirsch et al., 1989). Further evidence that sooty mangabeys are the natural primate reservoir of HIV-2 comes from the identical genomic organisation of HIV-2 and SIV_{SM} (Guyader et al., 1987; Hirsch et al., 1989), which both encode the *vpx* open reading frame (ORF) that is missing in HIV-1 and any other lentiviruses (Franchini et al., 1988; Yu et al., 1988).

Finally, this hypothesis is also supported by the geographical proximity and the close contacts existing between sooty mangabeys and humans in Western Africa, where sooty mangabeys are hunted for their meat and kept as pet animals (Hahn et al., 2000).

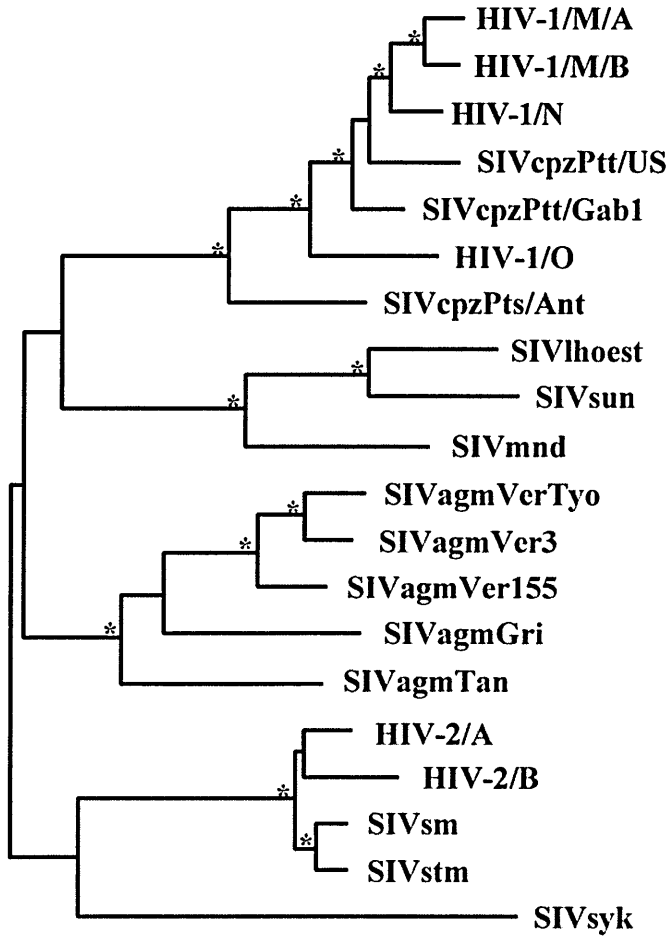


Figure 1.1: Evolutionary relationships of primate lentiviruses.
The tree was derived by maximum likelihood analysis of Pol protein sequences. Horizontal branch lengths are drawn to scale. Asterisks denote clades (to the right) supported in at least 80% of bootstrap replicates. Adapted from (Beer, 1999).

There are now eight reported HIV-2 groups or clades (formerly subtypes), A through H (Chen et al., 1997; Damond et al., 2004; Gao et al., 1994; Yamaguchi et al., 2000), each thought to have originated from a single cross-species transmission event (Chen et al., 1996; Gao et al., 1992; Hahn et al., 2000; Sharp et al., 1995). However, only group A and B were able to establish as human epidemics, with spread limited to Western Africa and countries such as Brazil, Portugal and France due to former colonisation, historical links and immigration pattern (Remy, 1998; Schim van der Loeff and Aaby, 1999; Smallman-Raynor and Cliff, 1991; Soriano et al., 2000). Groups C through H have only been isolated once, suggesting that they are primary zoonotic transfer events (Chen et al., 1997; Damond et al., 2004; Gao et al., 1994; Yamaguchi et al., 2000). In contrast to groups C-G, which have not been reported in symptomatic individuals, group H corresponds to a highly divergent HIV-2 isolate that is clearly pathogenic (Damond et al., 2004).

The most common recent ancestor of SIV in sooty mangabeys has been dated to the early 19th century and HIV-2 group A and B have been proposed to have entered the human population approximately in 1930 (Wertheim and Worobey, 2009).

- ***Evolution of HIV-1 from the SIV_{CPZ} lineage***

Serological cross-reactivity, similar genomic organisation and partial sequence homology between HIV-1 and SIV isolated from wild chimpanzees (*Pan troglodytes troglodytes*) in Gabon, SIV_{CPZ}, have been reported (Huet et al., 1990; Peeters et al., 1989), suggesting that HIV-1 may have evolved from the SIV_{CPZ} lineage.

Further analysis of different SIV_{CPZ} strains confirmed that HIV-1 clustered phylogenetically with SIV_{CPZ} (Gao et al., 1999; Keele et al., 2006; van Heuverswyn et

al., 2006) and that three cross-species transmission events gave rise to the major (M), new (N) and outlier (O) HIV-1 groups (see Figure 1.1).

While it seems well established now that chimpanzees are the natural reservoir of HIV-1 group M and N, the source of HIV-1 group O remains less clear (Keele et al., 2006). Interestingly, an HIV-1 group O-like virus was discovered in wild gorillas (*Gorilla gorilla*), SIV_{GOR}, and found to cluster with HIV-1 group O (Takehisa et al., 2009; van Heuverswyn et al., 2006). While the phylogenetic relationship between SIV_{CPZ}, SIV_{GOR} and HIV-1 suggests that *P. t. troglodytes* apes were the original reservoir of the viruses found in chimpanzees, gorillas and human, it is not known whether chimpanzees transmitted group-O like virus to gorillas and humans independently or whether gorillas served as an intermediary host for the human infection (Takehisa et al., 2009).

Despite having crossed the species barrier on three occasions, HIV-1 group M represents the only zoonosis that resulted in a pandemic. In fact, HIV-1 group N is only found in a few individuals in Cameroon (Simon et al., 1998; Yamaguchi et al., 2006) and HIV-1 group O seems to also localise around Cameroon (Peeters et al., 1997; Vanden Haesevelde et al., 1994), even though a few cases have been reported in Europe and USA, mainly in immigrants from central Africa (CDC, 1996; Loussert-Ajaka et al., 1995).

HIV-1 appears to have originated in west central Africa (Gabon, Cameroon, Equatorial Guinea and the Republic of Congo), where *P. t. troglodytes* and *G. gorilla* are found (Gao et al., 1999; Sharp et al., 2001; van Heuverswyn et al., 2006). Although it is difficult to precisely date the emergence of the HIV-1 lineage, the sequences of African HIV-1 group M isolates from the 1960's have been authenticated (Worobey et al., 2008; Zhu et al., 1998) and it has been proposed that HIV-1 group M emerged around 1930

(Hahn et al., 2000; Korber et al., 2000; Sharp et al., 2001; Sharp et al., 2000; Wertheim and Worobey, 2009).

1.2 The epidemiology of HIV

1.2.1 The global AIDS pandemic

According to the 2008 report on the global AIDS pandemic from the joint United Nations programme on HIV/AIDS (UNAIDS), there was an estimated 33 million [30 - 36 million] people worldwide living with HIV in 2007 (UNAIDS and World Health Organization, 2008). Although the annual number of new infections has decreased from an estimated 3 million in 2001 to 2.7 million in 2007, the overall number of people living with HIV has increased, partly due to the prolonged life expectancy of HIV patients on antiretroviral therapy (ART). The introduction and more widespread use of highly active ART (HAART) might also account for the reduction in the number of deaths due to AIDS from 2 million in 2001 to 1.7 million in 2007 (UNAIDS and World Health Organization, 2008).

1.2.2 The worldwide distribution of HIV

- ***The AIDS pandemic in Sub-Saharan Africa***

Even though the HIV pandemic has spread worldwide and is present on every continent (Figure 1.2 top panel), the African continent, and in particular Sub-Saharan Africa, bears a disproportionate share of the global burden of HIV.

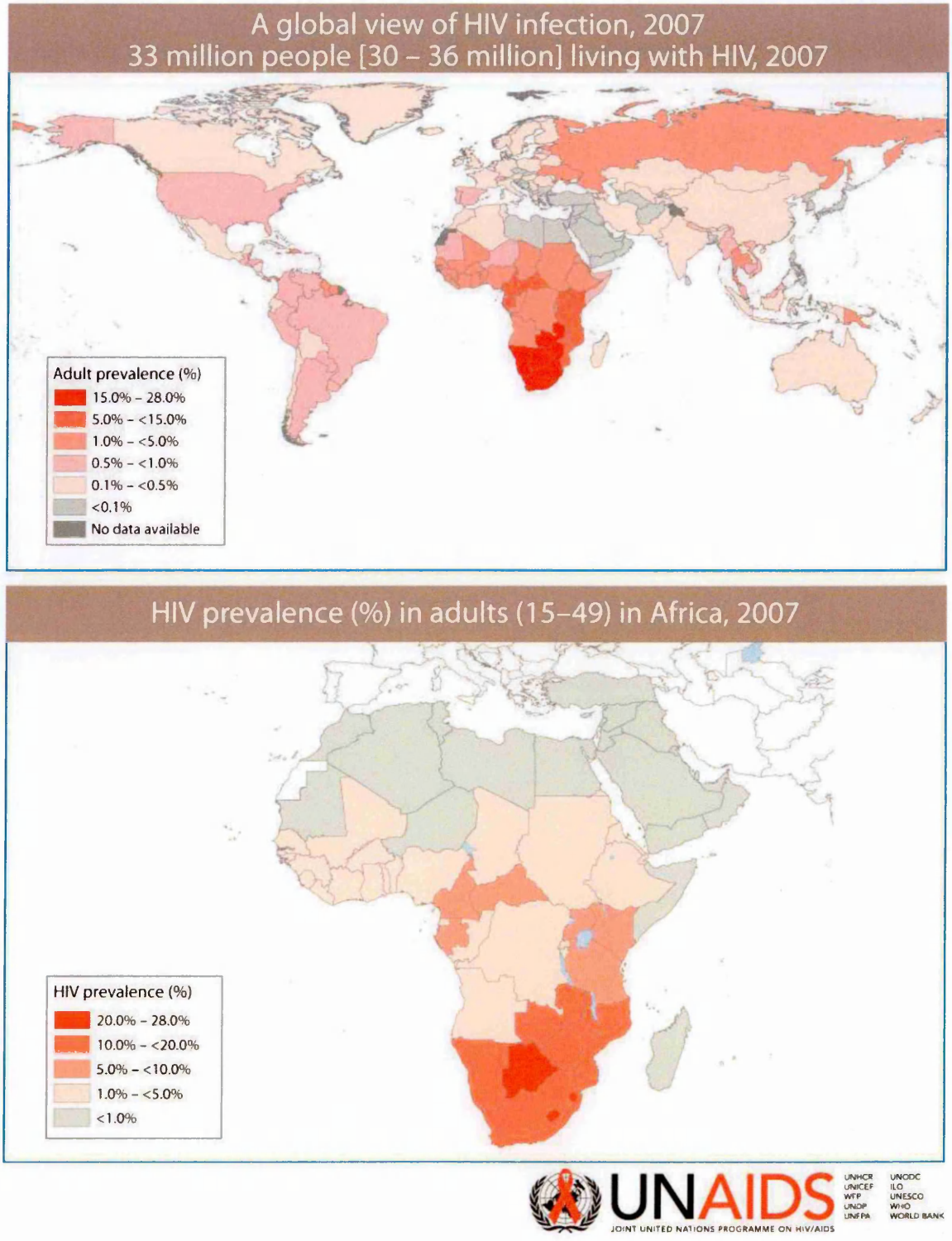


Figure 1.2: The global AIDS pandemic.
Percentage of HIV prevalence among adults worldwide (top panel; adults aged 15+) and in Africa (bottom panel; adults aged 15-49) in 2007 as reported by UNAIDS. Adapted from the 2008 report on the global AIDS pandemic from UNAIDS (UNAIDS and World Health Organization, 2008).

Indeed, Sub-Saharan Africa remains by far the most affected region, with an estimated 22 million people infected with HIV in 2007 and with HIV prevalence reaching almost 30 % in countries such as Mozambique (Figure 1.2 bottom panel). Sub-Saharan Africa is currently home to 67 % of all people living with HIV and 75 % of all AIDS deaths in 2007 occurred in this region (UNAIDS and World Health Organization, 2008).

- ***HIV in Asia***

Asia is the second most affected part of the world, with an estimated 5 million people living with HIV and 380,000 new cases in 2007 (UNAIDS and World Health Organization, 2008). South-East Asia (Thailand, Cambodia, Viet Nam and Indonesia) appears to have the highest HIV prevalence, between 1 and 2 %.

- ***HIV in Eastern Europe and Central Asia***

Eastern Europe and Central Asia account for approximately 4.5 % of all the HIV infections, with an estimated 1.5 million people living with the virus, mostly in Russia (69 %) and Ukraine (29 %). Approximately 4 % of the annual new infections and 2.9 % of AIDS deaths occur in this region (UNAIDS and World Health Organization, 2008).

- ***HIV in North America and Western Europe***

In Western Europe and North America, an estimated 81,000 people were newly infected with HIV in 2007, bringing the number of people living with HIV to 2 million. North America alone accounts for more than half of the cases, with 1.2 million infected people. Compared to other regions, relatively few people die of AIDS (31,000), most likely due to the widespread use and availability of HAART in those richer countries.

This also results in an increase in number of people living with the virus, as the life expectancy of patients is greatly augmented when HAART is used.

- ***HIV in Latin America and the Caribbean***

Latin America represents 5 % of HIV cases worldwide, with an estimated 1.7 million people living with the virus according to UNAIDS. 140,000 people became infected and 63,000 died because of AIDS in 2007.

In the Caribbean, 230,000 people live with HIV, three quarters of them in the Dominican Republic and Haiti. In 2007, an estimated 20,000 new infections and 14,000 deaths were reported for this region (UNAIDS and World Health Organization, 2008).

1.2.3 The HIV-2 epidemic

The emergence of HIV-2 occurred in Western Africa and to date, this region remains the only place where the virus is endemic (Schim van der Loeff and Aaby, 1999). Groups or subtypes A and B, the only HIV-2 groups that spread outside of Western Africa, account for the majority of HIV-2 infections (Norrgren et al., 1997; Schim van der Loeff and Aaby, 1999 and references therein).

In rural communities of Guinea-Bissau, HIV-2 prevalences of 8-10 % in the adult population (aged 15+) have been reported (Poulsen et al., 1989; Wilkins et al., 1993). Surrounding countries such as The Gambia, Senegal and Ivory Coast usually have lower prevalence rates, between 2 and 3 % (Remy, 1998), even though a prevalence of around 8 % was reported amongst commercial sex workers in Senegal and Ivory Coast (Ghys et al., 1997; Langley et al., 1996). This figure was reported to reach 28 % in The Gambia (reviewed in Reeves and Doms, 2002). HIV-2 has not spread to other regions of

Africa and the cases reported have been mainly sporadic ones (Schim van der Loeff and Aaby, 1999). Similarly, the HIV-2 prevalence in Europe and North America is very low, with the number of reported cases often below one hundred per country, and most cases can be traced back to Western Africa through the bias of population movements (Remy, 1998; Schim van der Loeff and Aaby, 1999). Due to their colonial past in Western Africa, Portugal and France are the two European countries with the highest HIV-2 prevalence (Smallman-Raynor and Cliff, 1991). In Portugal, HIV-2 accounts for 10-13 % of HIV infections and 4.5 % of AIDS cases (Soriano et al., 2000).

Although HIV-2 is endemic in Western Africa, the prevalence in this region is stable or in decline, while the rate of HIV-1 and dual HIV-1/HIV-2 infections has increased, suggesting that HIV-2 transmission (via the heterosexual route) is not as efficient as that of HIV-1 and that HIV-2 may be competed out by HIV-1 in the context of a dual infection (Anderson and May, 1996; Kanki et al., 1994; Schim van der Loeff and Aaby, 1999).

1.3 The pathogenesis of HIV

1.3.1 Course of infection with HIV

The typical course of an infection with HIV is presented in Figure 1.3. Following the initial exposure, patients experience an asymptomatic incubation period of approximately two to four weeks. In so-called 'typical progressors', the incubation period is followed by three phases: acute infection, clinical latency and AIDS, that span over eight to ten years (reviewed in Pantaleo and Fauci, 1996).

The acute phase of an HIV infection is characterised by a mononucleosis- or flu-like syndrome that appears three to six weeks after primary infection (Coffin et al., 1997).

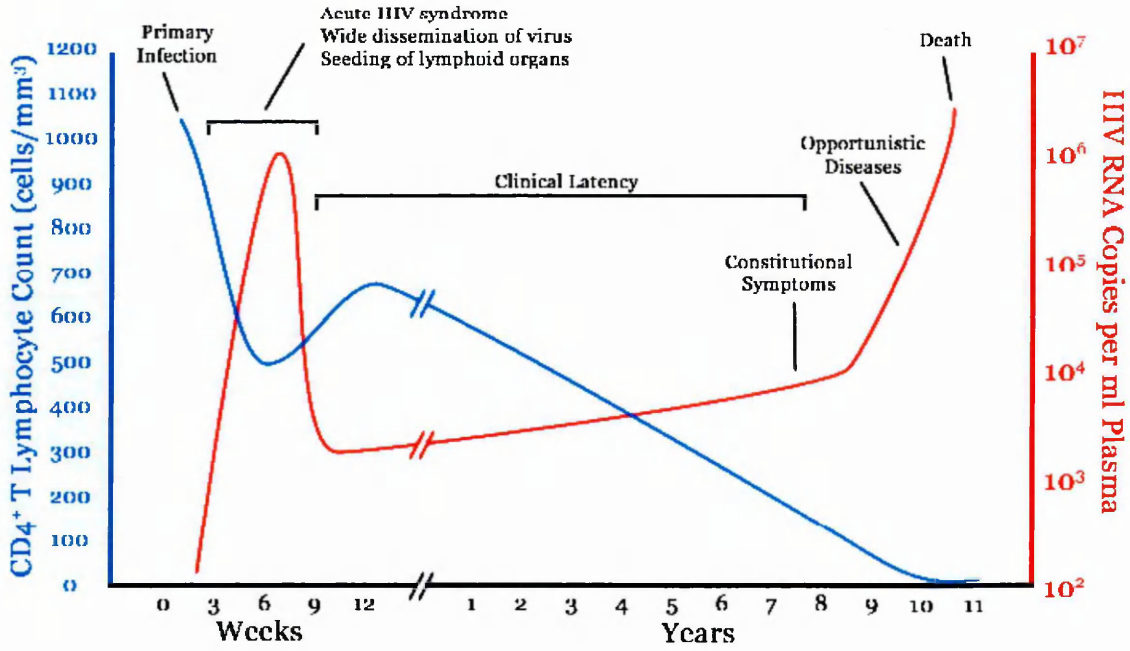


Figure 1.3: Typical course of an HIV infection.
Relationship between HIV genome copy number (viral load, red) and $CD4^+$ T-cell count (blue) over the typical course of an untreated HIV infection. The disease course, viral load and $CD4^+$ T-cell count may vary considerably from one patient to another (adapted from Pantaleo et al., 1993b).

Even though HIV targets $CD4^+$ cells such as helper T lymphocytes (Dalgleish et al., 1984; Klatzmann et al., 1984a; Klatzmann et al., 1984b), dendritic cells (DCs) are likely to be the first cells infected by the virus in the mucosa and are thought to play a key role in the dissemination of HIV in the infected host, and transmission to T-cells in particular (Granelli-Piperno et al., 1999; Rowland-Jones, 1999). During the acute phase, viral replication is rapid, with viral titres as high as 10^7 copies of HIV RNA per ml plasma, causing a steep decrease in the level of those $CD4^+$ T-cells in the peripheral blood in the first two to eight weeks following infection (see Figure 1.3; Gaines et al., 1990). The increased viremia is often associated with an activation of $CD8^+$ T-cells, which target the infected cells, and with antibody production. The efficacy of the $CD8^+$ T-cells response has been suggested to influence the progression of the disease (Pantaleo et al., 1997).

Following the acute phase of infection, typical progressors enter a long phase of clinical latency that can last for up to ten years. Viral replication remains active during latency and the count of $CD4^+$ T-cells decreases slowly (Embretson et al., 1993; Pantaleo et al., 1993a; Piatak et al., 1993). However, a $CD4^+$ T-cell count of 500 per µl or more is not usually associated with symptoms. Follicular dendritic cells (FDCs) appear to constitute a reservoir for the virus during the latent stage of infection and play an important role for viral persistence due to their close proximity to activated $CD4^+$ T-cells (Burton et al., 2002).

When the $CD4^+$ T lymphocyte count falls below 200 cells per µl, the disease progresses to AIDS and patients present an increased susceptibility to opportunistic infections.

Approximately 10-15 % of HIV infected patients, termed rapid progressors, present an unusually rapid progression to AIDS within two to three years of primary infection (for review, see Pantaleo and Fauci, 1996). On the other hand, a small percentage of patients

(5 %) do not show signs of disease progression over a prolonged period of time, up to twenty years, and remain asymptomatic. These individuals, called long-term non progressors (LTNPs), present stable $CD4^+$ T-cell counts (>500 cells per µl) for several years, increased $CD8^+$ T-cell count and often undetectable plasma viral load (Cao et al., 1995; Lifson et al., 1991; Pantaleo et al., 1995; Sheppard et al., 1993). Viral replication of HIV-1 and HIV-2 variants from LTNPs appears slower than that of viruses from normal progressors (Blaak et al., 1998; Blaak et al., 2006; Quinones-Mateu et al., 2000) and it was found that replication of HIV-2 variants from LTNPs replicated slower than HIV-1 from asymptomatic/LTNP individuals (Blaak et al., 2006). In LTNPs, the expression of inhibitory natural killer (NK) receptors on $CD8^+$ peripheral blood mononuclear cells (PBMCs) is similar to that of healthy individuals (Costa et al., 2003) and the NK cell cytotoxic response is reduced (Mavilio et al., 2003; O'Connor et al., 2007). In addition, neutralising antibodies titres are found to be higher than in normal progressors (Cao et al., 1995; Pantaleo et al., 1995). Finally, HIV-1 and HIV-2 variants isolated from LTNPs have been reported to carry large deletions in the *nef* (negative factor) gene, suggesting that the Nef protein plays a role in HIV pathogenesis (Blaak et al., 1998; Deacon et al., 1995; Kirchhoff et al., 1995; Rhodes et al., 2000; Switzer et al., 1998; Trible et al., 2007).

1.3.2 Differences in pathogenicity between HIV-1 and HIV-2

It has long been noted that HIV-2 is less virulent than HIV-1 and several studies have provided strong evidence supporting this difference in pathogenicity (reviewed in De Silva et al., 2008; Rowland-Jones and Whittle, 2007).

Firstly, the plasma viral load and replication rate of HIV-2 has been shown to be lower than for HIV-1 (Berry et al., 1998; Damond et al., 2002; Gottlieb et al., 2002; MacNeil

et al., 2007; Popper et al., 1999). However, amounts of total cellular and integrated proviral DNA are similar between HIV-1 and HIV-2, suggesting that poor proviral establishment cannot explain the lack of virulence of HIV-2 (MacNeil et al., 2007). Comparison between replication of HIV-1 and HIV-2 variants in long-term aviremic individuals showed a difference in the replication rate of the two viruses similar to that observed in normal progressors (Blaak et al., 2006).

The low plasma viral load may account for the low transmission rate observed with HIV-2 (reviewed in Jaffar et al., 2004; Schim van der Loeff and Aaby, 1999). In regions where both HIV-1 and HIV-2 are present, the prevalence of HIV-2 remains stable or decreases whereas HIV-1 prevalence increases, suggesting that the two viruses do not transmit at the same rate (Anderson and May, 1996; De Cock et al., 1993; Kanki et al., 1994). Mother-to-child HIV-2 transmission has also been shown to be a much rarer event than in HIV-1 infections and HIV-1/HIV-2 dually infected mothers seemed to transmit HIV-1 alone in the majority of cases (Adjorlolo-Johnson et al., 1994; Ekpini et al., 1997; O'Donovan et al., 2000).

An enhanced immune response is also a characteristic of HIV-2 infection. Long-term non progression is more common in HIV-2 than HIV-1 infections (Berry et al., 2002; Poulsen et al., 1989) and the rate of $CD4^+$ T-cell decline in HIV-2 infected individuals is slower compared to HIV-1 infection (Jaffar et al., 1997). This higher $CD4^+$ T-cell count is associated with a greater NK cell cytolytic response and broader cytokine secretion compared to HIV-1 infection (Nuvor et al., 2006).

Strong Gag-specific $CD4^+$ T-cell response and $CD8^+$ cytotoxic T lymphocyte (CTL) response have been proposed to further help control viremia in HIV-2 infected patients (Duvall et al., 2006; Jennes et al., 2008; Leligdowicz et al., 2007) and HIV-2-specific T-

cells have been shown to maintain a polyfunctional profile (Duvall et al., 2008), constituting a hallmark for non-progressive HIV-2 infection.

In addition, the neutralising antibody response has been shown to be different between HIV-1 and HIV-2 infections. HIV-2 infection triggers a broad, low-magnitude neutralising response whereas HIV-1 is characterised by a narrow, high-magnitude response (Rodriguez et al., 2007).

Interestingly, dendritic cells were shown to be less susceptible to HIV-2 than to HIV-1 infection (Duvall et al., 2007). Since DCs have been proposed to be the 'port of entry' of HIV upon primary infection and to be implicated in the dissemination of HIV, this might partly account for the lower pathogenicity of HIV-2.

Finally, immune activation has been implicated in the pathogenicity of HIV (De Silva et al., 2008 and references therein). Contrary to HIV-infected human and SIV-infected Rhesus macaques, the observation that SIV-infected sooty mangabeys do not develop AIDS despite high viremia and low level of specific $CD8^+$ T-cell response led to the suggestion that immune activation may play an important role in the pathogenicity of HIV (Silvestri et al., 2003). Indeed, maintenance and regeneration of the T lymphocyte population and lower levels of aberrant immune activation and apoptosis were observed compared to pathogenic HIV and SIV infections. Accordingly, a recent study of SIV_{SM} in its natural host has suggested that the killing of the $CD4^+$ T-cells by HIV might not represent the main determinant of disease progression (Gordon et al., 2008). Interestingly, the HIV-2 and SIV_{SM} Nef proteins were shown to downregulate the T-cell receptor (TCR)-CD3 complex on infected cells, preventing overactivation of infected cells and activation-induced cell death. This Nef-mediated suppression of immune activation present in HIV-2 and SIV_{SM} was lost in the HIV-1 lineage (Schindler et al., 2006) and may contribute to the difference in virulence between HIV-1 and HIV-2.

Supporting this hypothesis, HIV-1 variants isolated from LTNPs possessed large deletions in the *nef* gene (Blaak et al., 1998; Kirchhoff et al., 1995; Rhodes et al., 2000; Trible et al., 2007). Similarly, truncations in the *nef* ORF appear characteristic of attenuated HIV-2 infections as these were found in 14 % of asymptomatic HIV-2 infections but in only 4 % of normal HIV-2 infected progressors (Switzer et al., 1998).

1.3.3 Interaction with the host

There has been increasing evidence that host species have developed certain innate immunity mechanisms to counteract retroviral replication. For example, proteins from the family of human apolipoproteins, most notably APOBEC3G (apolipoprotein B mRNA-editing enzyme catalytic subunit 3G), were shown to restrict HIV-1 and HIV-2 replication by using both a deaminase-dependent and -independent mechanism (reviewed in Holmes et al., 2007). This restriction was shown to be species-specific (Bogerd et al., 2004; Mariani et al., 2003) and to be counteracted by the Vif (virion infectivity factor) protein of HIV, which prevents the packaging of APOBEC3G into nascent budding virions and triggers its proteasomal degradation (Sheehy et al., 2003). Another restriction factor, Tripartite motif protein 5α (TRIM5α), was later identified and shown to restrict HIV replication in RhMac cells (reviewed in Huthoff and Towers, 2008). Similarly, SIV replication was found to be restricted by human TRIM5α. The TRIM5α proteins act by binding to the retroviral capsid (CA) and interfering with reverse transcription as well as later stages of viral replication. It was recently reported that a specific residue in HIV-2 CA (proline 119) determines sensitivity to cynomolgus monkey and human TRIM5α (Song et al., 2007a). Interestingly, a study in the HIV-2 community cohort Caio in Guinea-Bissau has revealed that HIV-2 CA encoding proline residues at three key positions, including at position 119, were found more frequently in

individuals with low viral load (Onyango et al., unpublished data reported in De Silva et al., 2008). This so-called p26PPP capsid seemed to be less stable and the PPP-encoding viruses appeared to generate stronger T-cell response compared to viruses lacking the proline residues. It was suggested that the presence of prolines at these positions might promote a better exposure of CA to TRIM5α-mediated restriction, leading to a better control of viral replication by the host. Supporting this hypothesis, non-PPP-encoding viruses were found to be enriched in HIV-2 infected patients progressing to AIDS. This implies that specific capsid sequences could influence the progression of the disease and therefore represent an interesting tool to help predict and monitor disease progression. Conversely, the presence of PPP in the capsid of HIV-2 isolated from AIDS patients such as HIV-2_{ROD} shows that this factor alone cannot be considered a marker of low viral load and slow disease progression.

It is interesting to note that TRIM5α is strongly induced by interferon (IFN) and that HIV-2 has been shown to be restricted by high levels of human TRIM5α (Ylinen et al., 2005). Another IFN-induced tripartite motif protein, TRIM22, was also shown to restrict HIV-1 replication by inhibiting the budding of Gag virus-like particles (VLPs) and affecting Gag trafficking (Barr et al., 2008).

More recently, a protein named tetherin (also known as B cell stromal factor 2 [Bst-2] or CD317) was shown to block HIV-1 particle release (Neil et al., 2008; Van Damme et al., 2008). Tetherin is IFN-induced, similarly to TRIM5α and TRIM22, and its antiviral activity was shown to be blocked by the viral protein u (Vpu) in the case of HIV-1 and Env for HIV-2.

1.4 The HIV particle

Like all retroviruses studied to date, HIV-1 and HIV-2 virions contain two copies of their single-stranded RNA genome, non-covalently linked near their 5′ end (Coffin et al., 1997). The virus is released from the cell in an immature form that undergoes maturation via proteolytic processing of the Gag and Gag-Pol polyproteins by the virally encoded protease. The mature virion contains a cone-shaped condensed core composed of the capsid protein and containing the dimeric genome coated by the nucleocapsid protein as well as the viral enzymes (reverse transcriptase, integrase, protease) necessary for HIV replication in the target cell (see Figure 1.4). Enveloping the capsid, the cell-derived lipid membrane is coated with the matrix protein on the inside and the envelope glycoproteins on the exposed surface.

1.4.1 HIV genome organisation

Despite sequence homology of only 50 to 60 %, HIV-1 and HIV-2 share a relatively similar genomic organisation (Alizon et al., 1984; Clavel et al., 1986b; Franchini et al., 1987; Guyader et al., 1987).

The genome of the HIV-2_{ROD} isolate, shown in Figure 1.5, is a 9671 nucleotides (nt) long positive-sense single strand RNA and comprises three major ORFs, *gag*, *pol* and *env*, as well as six accessory genes, *vif*, *nef*, *tat* (*trans*-activator of transcription), *rev* (regulator of viral expression), *vpr* (viral protein r) and *vpx* (viral protein x; Guyader et al., 1987).

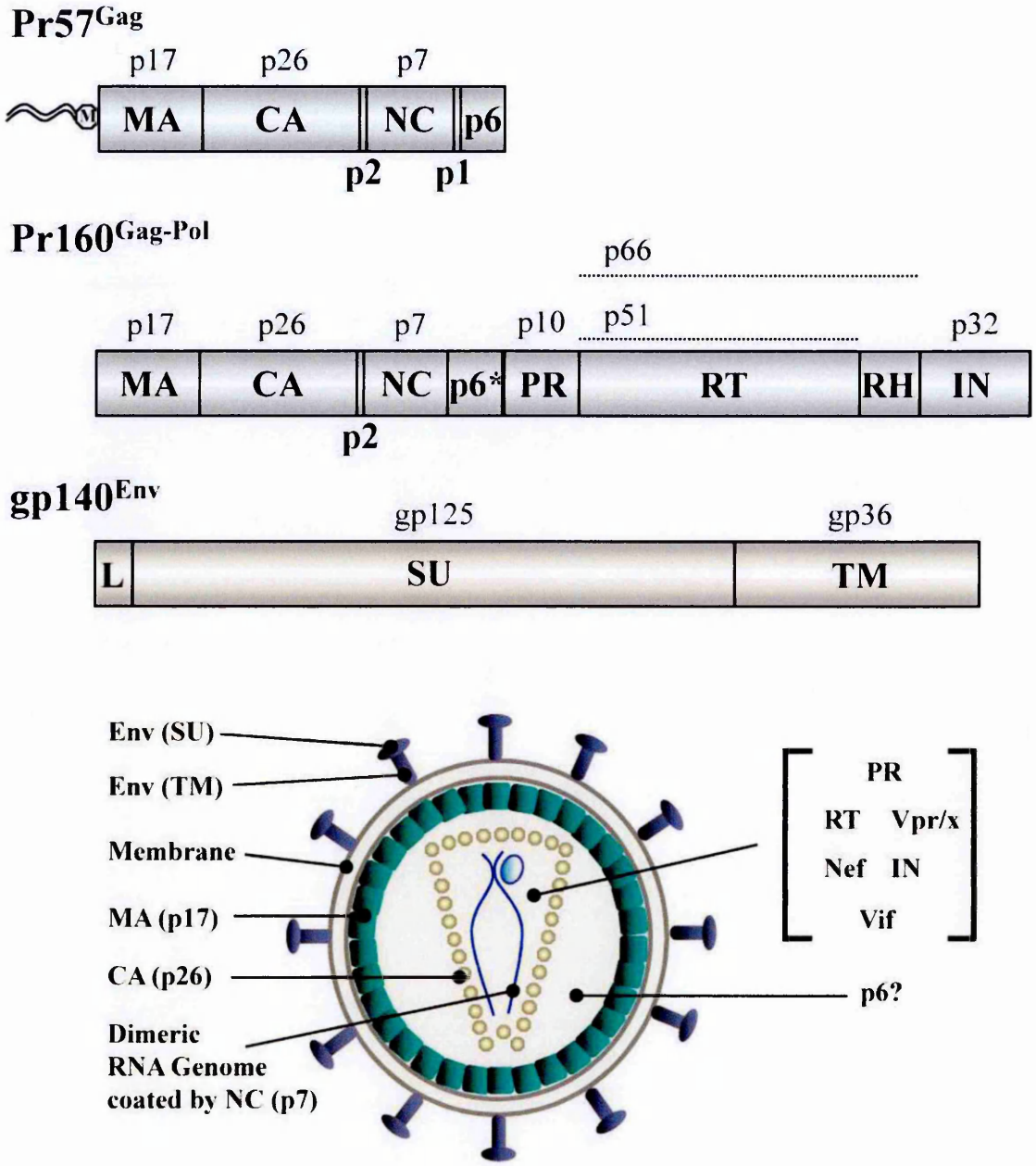


Figure 1.4: Major components of the HIV-2 virion.
Schematic representation of the Gag, Gag-Pol and Env precursors and of a mature HIV-2 particle. The name and molecular weight in kDa of the domains released following proteolytic cleavage of each precursor are indicated. Abbreviations are as defined in the text. Figure courtesy of S.D. Griffin.

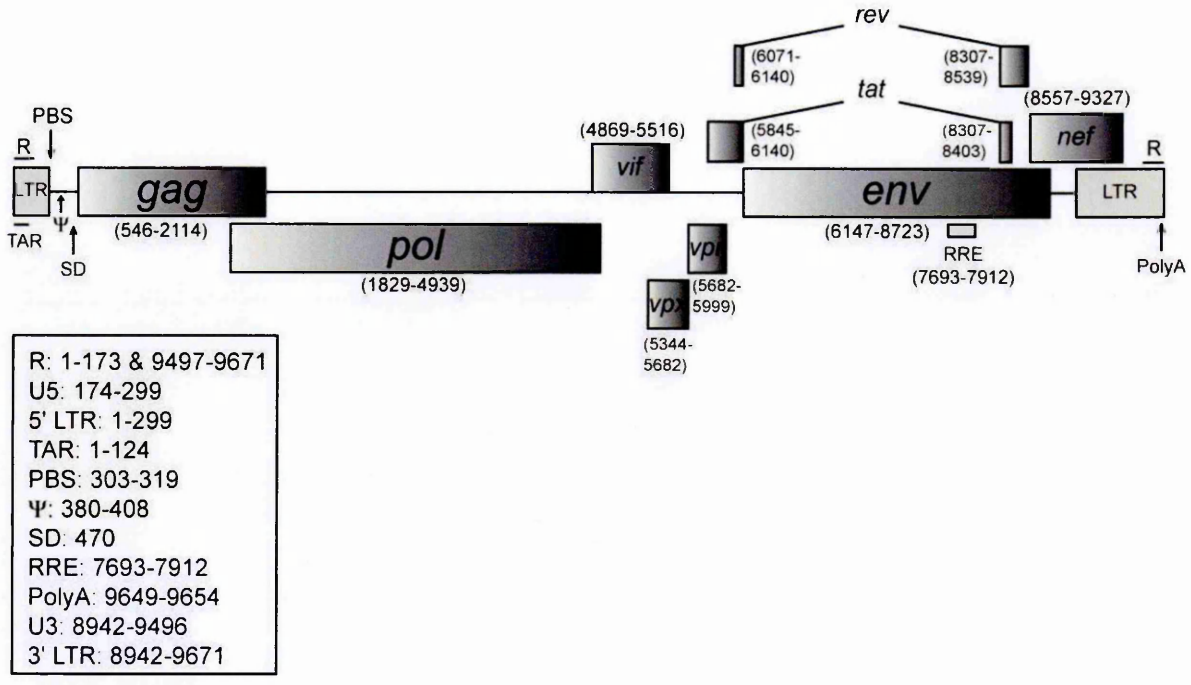


Figure 1.5: HIV-2 genome organisation.

The HIV-2 genome is a single-stranded positive-sense RNA of 9671 nt. The viral genome contains two long terminal repeats (LTRs) and several ORFs encoding the Gag, Pol, Env, Vif, Vpx, Vpr, Tat, Rev and Nef proteins. The 5′ and 3′ untranslated regions comprise the regulatory elements: TAR, *trans*-activation responsive; PBS, primer binding site; SD, splice donor; Ψ, packaging signal; PolyA, poly(A) signal. The Rev responsive element (RRE) is located within the *env* ORF. The 5′ LTR includes a repeat (R) and a unique 5′ (U5) region. The 3′ LTR includes a unique 3′ (U3) and R region. The length and position of each element is indicated in the box. Figure is to scale.

As shown in Figure 1.4, the *gag* gene encodes the structural proteins matrix (MA), capsid (CA) and nucleocapsid (NC), two spacer peptides p1 and p2 and the late domain p6. The *pol* ORF is divided into *pro* and *pol* and encodes the viral enzymes. *Pro* encodes the viral protease (PR) and *Pol* encodes the RNA-dependent DNA polymerase or reverse transcriptase (RT) and the integrase (IN). The reverse transcriptase also exerts an RNase H (RH; DNA-RNA heteroduplex dependent Ribonuclease) activity. The *env* gene encodes the Envelope precursor glycoprotein (gp), gp140 and gp160 in HIV-2 and HIV-1 respectively. Cleavage of the precursor by a host protease gives rise to the surface (SU) and transmembrane (TM) glycoproteins, gp125 and gp120 and gp36 and gp40 for HIV-2 and HIV-1 respectively (Clavel et al., 1986a; Freed et al., 1989; McCune et al., 1988).

The six accessory genes are located between the *pol* ORF and the U3 region and encode proteins involved in transcription (Tat), viral infectivity (Vif, Vpr and Vpx), mRNA export (Rev) and downregulation of cell surface CD4 and MHC (major histocompatibility complex) class I (Nef; reviewed by Cullen, 1998). All six genes are present in HIV-1 as well as HIV-2 except for *vpx*, which is only present in HIV-2 and SIVs isolated from most species and is thought to have arisen from a duplication of an ancestor *vpr* gene (Tristem et al., 1992). In HIV-1, the *vpx* gene is replaced by *vpu*, a unique feature of the SIV_{CPZ}/HIV-1 lineage. A detailed description of the function of each protein is presented is section 1.4.2.

Located at the extremities of the RNA genome, the untranslated regions (UTR) contain regulatory elements important for several stage of the virus life cycle such as transcription, splicing, reverse transcription, genome dimerisation and encapsidation. In particular, the 5′ leader RNA comprises the *tat*-responsive element TAR, the primer binding site (PBS), the major splice donor (SD), the packaging signal (Psi or Ψ) and the

Polyadenylation (polyA) signal, which is also present at the 3′ end of the genome due to its location within the repeat (R) region. Finally, the Rev-responsive element (RRE) is located within the *env* ORF.

At the 5′ and 3′ ends of the genome lie the long terminal repeats (LTR) composed of a repeat (R) and a unique (U5 and U3, respectively) region. The unique regions are duplicated during reverse transcription when the RT enzyme jumps from the 5′ to the 3′ end of the genomic RNA template (see section 1.5.2).

Like the cellular messenger RNAs (mRNAs), the RNA genome and the viral mRNAs are capped at their 5′ end and polyadenylated at their 3′ end.

1.4.2 HIV proteins

- ***The structural protein Gag***

The Gag polyprotein is translated from the full-length unspliced genomic RNA and has a molecular weight (MW) of approximately 57 kilodaltons (kDa), although two other HIV-2 Gag isoforms translated from downstream initiation codons have been described (p50 and p44; Herbreteau et al., 2005).

The polyprotein is cleaved by the viral protease during particle maturation to produce the matrix (p17), capsid (p26) and nucleocapsid (p7) proteins as well as the late domain p6 and two spacer peptides p1 and p2 (see Figure 1.4). The five cleavage sites were found to be processed at different rates and in the following order: NC/p2 > MA/CA ≈ p1/p6 > NC/p1 > CA/p2 (for review see Ganser-Pornillos et al., 2008 and references therein). The specific sequence and rate of cleavage have been shown to be important for particle morphogenesis and infectivity (Erickson-Viitanen et al., 1989; Pettit et al., 1994; Tritch et al., 1991).

Matrix

Located at the amino (N)-terminus of the Gag (p57) polyprotein, the HIV-2 matrix protein is 134 amino acids (aa) long, with a MW of approximately 17 kDa. A 14-carbon fatty acid (myristate) is cotranslationally added to a glycine at the N-terminus of MA, suggesting that it is associated to lipid bilayers (Jacobs et al., 1989; Rein et al., 1986). It was therefore suggested that matrix could play a role in targeting the Gag polyprotein to the plasma membrane during the assembly process and myristylation was indeed shown to be absolutely required for retroviral particle assembly (Bryant and Ratner, 1990; Rein et al., 1986; Wang and Barklis, 1993). In addition, myristylation of MA has been proposed to influence MA multimerisation during assembly (Bouamr et al., 2003).

MA was also reported to participate in the recruitment of the viral envelope glycoprotein into nascent particles (Dorfman et al., 1994b; Freed and Martin, 1995; Yu et al., 1992). Most of MA, apart from its carboxy (C)-terminus, was found to be required for Env incorporation (Dorfman et al., 1994b) and it appeared that minor changes in the conformation of MA were sufficient to abrogate the interaction between the two proteins (Yu et al., 1992). N-terminal residues of HIV-1 MA, together with residues at the C-terminus of the gp41 cytoplasmic tail, were shown to be involved in the Env-MA interaction (Freed and Martin, 1996; Murakami and Freed, 2000).

In addition to its role in particle morphogenesis, MA has been implicated in an early step post-infection. A small amount of mature phosphorylated HIV-2 MA can be detected inside the viral core (Gallay et al., 1995b) and MA was reported to associate with components of the pre-integration complex (PIC; Bukrinsky et al., 1993b). HIV-1 MA was shown to promote the nuclear import of the PIC during infection of non-dividing cells (Bukrinsky et al., 1993a). To do so, it was proposed that a tyrosine residue at the C-terminus of HIV-1 MA was phosphorylated by a cellular protein kinase,

resulting in the interaction of MA with the integrase (Gallay et al., 1995a; Gallay et al., 1995b). However, it was subsequently demonstrated that the infectivity defect of the MA mutants was only observed in the absence of a functional Vpr (Heinzinger et al., 1994; von Schwedler et al., 1994). Furthermore, the MA-IN interaction was not confirmed (Fouchier et al., 1997) and the role of MA in early post-entry steps therefore remains unclear.

While MA plays a major role in particle assembly through membrane association and envelope incorporation, it does not appear to be involved in HIV genome encapsidation (Lee and Linial, 1994; Poon et al., 1998). However, in other retroviruses, MA has been linked with genome dimerisation and RSV MA mutants were found to display a dimerisation defect (Parent et al., 2000).

Capsid

CA (p26) is the most abundant protein in the virion. CA is composed of two domains; the N-terminal 'core' domain (NTD; residues 1-145), involved in virion maturation and cyclophilin A (cypA) binding, and the C-terminal 'dimerisation' domain (CTD; residues 151-231), which functions in Gag-Gag interaction and particle assembly (reviewed by Freed, 1998). The C-terminal region of CA comprises a stretch of 20 conserved amino acids, termed the major homology region (MHR), which is essential for viral replication (Mammano et al., 1994).

The cellular peptidyl prolyl isomerase cyclophilin A was shown to specifically interact with the NTD of HIV-1 CA, and the immunosuppressive drug cyclosporine A and its non-immunosuppressive analogues prevented this interaction, blocking the incorporation of CypA into HIV-1 virions and impairing HIV-1 replication (Franke and Luban, 1996; Franke et al., 1994b; Gamble et al., 1996; Luban et al., 1993; Thali et al.,

1994). CypA has also been shown to interact with the capsid of incoming virions in the target cells (Hatziioannou et al., 2005). The CA-CypA interaction was proposed to be important at an early stage of the virus life cycle (Thali et al., 1994). CypA was postulated to play a role in the uptake of HIV in PBMCs and T-cells (Sherry et al., 1998) and viral cDNA synthesis was found to be reduced in the absence of CypA incorporation (Braaten et al., 1996). A study that analysed the effect of CypA on viral replication independently of cell attachment suggested that CypA plays distinct roles in entry and post-entry events of HIV-1 replication (Saphire et al., 2002). Since cyclophilins such as CypA are involved in the regulation of protein folding, incorporation of CypA into HIV-1 virions may be required for the uncoating of the viral nucleoprotein complex following viral entry. The crystal structure of the NTD of CA in complex with CypA revealed that CypA could weaken the association between the CA strips that composed the capsid lattice, therefore promoting the disassembly of the viral core (Gamble et al., 1996).

In addition to its role in early post-entry event, CA plays a crucial role during particle assembly (reviewed in Ganser-Pornillos et al., 2008). Mutations in the CA CTD were shown to impair viral replication and reduce particle production, suggesting that the CTD is required for particle assembly (Dorfman et al., 1994a; Joshi et al., 2006; Jowett et al., 1992; Zhang et al., 1996). Gag multimerisation was affected by CTD mutations (Burniston et al., 1999; Franke et al., 1994a) and structural data supported a role in viral assembly (Gamble et al., 1997). Further mutational analysis of the CTD dimer interface, as defined by crystallography studies (Gamble et al., 1997), confirmed that CA dimerisation was important for particle assembly (Ganser-Pornillos et al., 2004; von Schwedler et al., 2003a). Mutations of the MHR also resulted in a severe defect in particle production (Mammano et al., 1994) and Gag multimerisation (Chang et al.,

2007). In contrast, mutations of the NTD of CA blocked viral replication but did not affect particle assembly and release. Interestingly, CA NTD mutations impaired particle morphogenesis, especially mature capsid formation, as the virions produced contained circular or aberrant cores instead of cone-shaped cores (Dorfman et al., 1994a; Scholz et al., 2005; Tang et al., 2007; von Schwedler et al., 1998; von Schwedler et al., 2003a). Consequently, a model was proposed for the role of CA NTD in particle assembly (von Schwedler et al., 1998). Following proteolytic cleavage of the MA-CA junction, the NTD of CA is refolded into a beta-hairpin/helix structure, creating a new CA-CA dimer interface that is essential for the formation of a condensed conical core.

It has also been suggested that CA may be required for the incorporation of the Gag-Pol precursor into virions, a step that is necessary for the recruitment of the RT, IN and PR enzymes into the budding virions (Huang and Martin, 1997; Srinivasakumar et al., 1995). In particular, the MHR and CTD of CA were found to be critical for the packaging of the Gag-Pol precursor, although other regions of CA are involved.

Finally, HIV-1 CA has also been implicated in RNA dimerisation (Kafaie et al., 2009). Disruption of helix 1, at the N-terminus of CA, was shown to trigger a dimerisation defect similar to that observed in protease-deficient viruses.

Nucleocapsid

NC is a small basic protein that coats the genomic RNA in virions. During proteolysis, HIV NC can be found in three forms, p15, p9 and p7, which correspond to the precursors NC-p1-p6, NC-p1 and the mature NC, respectively. The NC domain of Gag p55, the proteolytic cleavage intermediates p15 and p9, and the fully cleaved NC p7 protein all possess nucleic acid binding activity (Berkowitz et al., 1993; Cruceanu et al., 2006).

NC contains two conserved cysteine-histidine boxes (C-X_2-C-X_4-H-X_4-C; CCHC) called zinc fingers motifs (Covey, 1986; South et al., 1990). The CCHC motifs of retroviral NC proteins were found to bind zinc (Green and Berg, 1989; South et al., 1990) and mutations that inhibited zinc binding led to the production of non-infectious particles (Meric and Goff, 1989; Meric et al., 1988), suggesting that the zinc binding domain of NC plays an important role in viral infectivity. In HIV-1, the two zinc fingers are not functionally equivalent. The first (N-terminal) CCHC motif was shown to be critical for RNA binding and encapsidation as well as viral replication, whereas the second CCHC motif appeared dispensable for RNA packaging (Dorfman et al., 1993; Gorelick et al., 1993; Schwartz et al., 1997). The context of the zinc finger also appeared to be important as inverting the position of the two motifs led to a strong defect. The situation is different in HIV-2, where both CCHC motifs are required for RNA binding (Matsui et al., 2009).

NC is involved in many crucial steps of the virus life cycle (reviewed by Thomas and Gorelick, 2008). For example, NC is a key element in the selection of the genomic RNA during packaging and it facilitates the stabilisation of the genomic RNA dimer during particle maturation, as described in sections 1.6.3 and 1.7.2 respectively. In addition, many mutations in NC result in defects in virus assembly and release (Dorfman et al., 1993; Grigorov et al., 2007; Poon et al., 1996), suggesting that NC plays a role in particle assembly, and it has been proposed that NC participates in Gag multimerisation (Burniston et al., 1999; Cimarelli et al., 2000; Franke et al., 1994a). Finally, NC has been shown to promote the placement of the $tRNA^{lys3}$ primer onto the PBS (Feng et al., 1999; Hargittai et al., 2004; Huang et al., 1998; Prats et al., 1988) and to be required for efficient reverse transcription and integration of the PIC into the genome of the host cell (Thomas and Gorelick, 2008). For example, NC has been implicated in the initiation of

the minus-strand synthesis from the $tRNA^{lys}$ (Li et al., 1996), strand transfer and specific plus-strand priming (Guo et al., 1997; Post et al., 2009).

The late domain p6

As its name suggests, the late domain p6 plays a major role late in the virus life cycle by allowing the release of the budding particles from the infected cells in a membrane fission event (reviewed in Freed, 1998; Fujii et al., 2007; Gottlinger, 2001; Martin-Serrano, 2007). Indeed, HIV-1 p6 mutant viruses appear to be tethered to the plasma membrane and therefore accumulate at the cell surface (Gottlinger et al., 1991). The release function of p6 is mediated by a PTAP amino acid motif, conserved amongst lentiviruses (Gottlinger et al., 1991; Huang et al., 1995), or a conserved PPxY motif in the case of the late 'L' domain of oncoretroviruses such as RSV (Wills et al., 1994). In HIV, p6 mediates the release of the nascent virions by interacting with components of the ESCRT (endosomal sorting complexes required for transport) machinery (see Fujii et al., 2007; and Martin-Serrano, 2007 for review). In particular, the ESCRT-I component Tsg101 (tumor susceptibility gene 101) has been shown to interact with the PTAP motif of HIV-1 p6, promoting HIV-1 release (Garrus et al., 2001; Martin-Serrano et al., 2001). In addition, HIV-1 p6 and EIAV p9 interact with a cellular protein called AIP1/Alix (ALG-2-interacting protein 1) through a conserved YPx_nL motif (Strack et al., 2003).

In addition to its role in particle release, the p6 late domain was shown to promote the encapsidation of Vpr in viral particles, mainly through the binding of Vpr to a conserved $(Lxx)_4$ sequence (Jenkins et al., 2001; Kondo and Gottlinger, 1996; Lu et al., 1995; Paxton et al., 1993) and an FxFG motif (Zhu et al., 2004).

- ***The polymerase and protease***

The viral protease and polymerase are translated as part of the Gag-Pol precursor from the unspliced genomic RNA following a -1 frameshifting event during Gag translation (Jacks et al., 1988). Ribosomal frameshifting occurs at the site of a slippery sequence followed by a hairpin structure. The role of the latter is thought to be to slow the ribosome down and allow the -1 frameshift (reviewed in Brierley and Dos Ramos, 2006). A 7 nt spacer is located between the slippery sequence and the hairpin structure of HIV-1 but is absent in HIV-2. In HIV-1, the ratio of Gag/Gag-Pol is 20:1, corresponding to a frameshifting efficiency of 5 %, and has been shown to be critical for viral infectivity, genome dimerisation and particle assembly (Karacostas et al., 1993; Shehu-Xhilaga et al., 2001a).

The Gag-Pol precursor is processed by the viral protease PR into MA, CA, p2, NC, p6* (a modified p6 due to the frameshift), PR (p10), RT (p51 and p66) and IN (p32), as depicted in Figure 1.4. The RT enzyme comprises an RNA-dependent DNA polymerase and an RNase H activity and accomplishes the reverse transcription of the RNA genome into the proviral cDNA. Finally, the integrase triggers the integration of the proviral cDNA into the genome of the host cell.

- ***The envelope glycoprotein***

The HIV-2 envelope glycoprotein is translated from a singly spliced mRNA of approximately 4.3 kb. The Env precursor (gp140) is cleaved by a cellular protease into two subunits, the surface glycoprotein gp125 and the transmembrane glycoprotein gp36 and contains a leader (L) signal at its N-terminus (see Figure 1.4). Env mediates viral entry by binding of gp125 to the CD4 receptor and CC-chemokine receptor 5 (CCR5) or

CXC-chemokine receptor 4 (CXCR4) co-receptor. The viral envelope subsequently fused with the plasma membrane due to the fusogenic activity of gp36.

The HIV-2 Env is also involved in the release of HIV-2 particles from infected cells, an activity that has been named enhancement of virus release (Bour et al., 2003; Bour et al., 1996; Ritter et al., 1996). This function of HIV-2 Env mimics the activity of Vpu in HIV-1, which was also shown to enhance virus release in a cell-type dependent manner (Strebel et al., 1989; Terwilliger et al., 1989), and is assumed by two distinct motifs located in the cytoplasmic tail and the ectodomain of Env (Abada et al., 2005). It was suggested that a cellular factor was involved in the inhibition of HIV-1 release and that Vpu blocked this antiviral activity (Neil et al., 2007). Indeed, HIV-1 was shown to be restricted by a host protein termed tetherin/Bst-2, which prevents the release of budding virions from the plasma membrane of infected cells (Neil et al., 2008; Van Damme et al., 2008). In HIV-1, the accessory protein Vpu counteracts the antiviral activity of tetherin and enhances particle release. In HIV-2, this function is assumed by the Env glycoprotein, using a mechanism as yet unknown but likely to involved a cellular co-factor (Stuart Neil, personal communication), even though there have been some evidence that HIV-2 Env is able to down-regulate surface expression of Bst-2 (John Guatelli, personal communication) and to re-localise Bst-2 to a cytoplasmic compartment (Paula Cannon, personal communication).

- ***Accessory proteins***

HIV-1 and -2 both encode six accessory proteins. Tat, Rev and Nef are translated from multiply spliced viral mRNAs of approximately 2 kb, while Vif, Vpu/Vpx and Vpr are synthesised from singly spliced mRNAs of approximately 4 kb.

Fully spliced mRNA transcripts can be exported out of the nucleus by the cellular machinery and are therefore translated relatively early in the virus life cycle. On the other hand, the nuclear export of singly spliced mRNAs, similarly to that of the unspliced genomic RNA, is inhibited by the host cell machinery in charge of preventing the export and translation of incompletely spliced cellular mRNAs. As a result, HIV-1 and HIV-2 have developed an independent way to export intron-containing mRNAs, using the Rev protein (Cullen, 1998; Malim et al., 1989a; Malim et al., 1989b). Because it contains both a nuclear localisation and a nuclear export signal (NLS and NES, respectively), Rev is able to shuttle in and out of the nucleus and triggers the nuclear export of incompletely spliced viral RNAs by binding to an RNA target located in the *env* ORF, the Rev-responsive element or RRE (Malim and Cullen, 1991; Malim et al., 1990; Tiley et al., 1992). To mediate nuclear export, Rev interacts with the cellular Crm1 (chromosome region maintenance 1) protein, a nucleocytoplasmic shuttle protein which is able to interact with the Rev NES and cellular nucleoporins (Bogerd et al., 1998). Interestingly, the ATP-dependent DEAD box RNA helicase DDX3 was shown to be required for the Crm1/Rev-mediated nuclear export activity (Yedavalli et al., 2004). In addition to its crucial role in the regulation of viral protein expression, Rev has been proposed to affect RNA packaging and protein translation in HIV-1 (reviewed in Groom et al., 2009).

Tat is a *trans*-activator of HIV-1 and HIV-2 LTR-driven gene expression (Cullen, 1986; Dayton et al., 1986; Guyader et al., 1987), which regulates the transcription of the provirus by binding to the *trans*-activation responsive element TAR located downstream of the promoter region, in the 5′ leader RNA (Fenrick et al., 1989; Jakobovits et al., 1988). The structure of TAR, and in particular the presence of an exposed loop and a side bulge, were shown to be important for *trans*-activation (Berkhout et al., 1990;

Berkhout and Jeang, 1989; Feng and Holland, 1988). Tat-mediated *trans*-activation of HIV gene expression was suggested to involve the phosphorylation of the RNA polymerase II (Pol II) CTD. It has been demonstrated that Tat forms a complex with a cellular protein termed cyclin T1 and that this Tat-cyclin T complex binds to the TAR stem-loop with high affinity. Tat was also shown to immunoprecipitate with a cyclin-dependent kinase, CDK9 (Zhu et al., 1997), which appears to be recruited by cyclin T1 (Wei et al., 1998). CDK9 and cyclin T1 form the positive transcription elongation factor b (P-TEFb). Recruitment of CDK9 leads to the phosphorylation of Pol II CTD and enhancement of its processivity (reviewed by Cullen, 1998).

The negative factor Nef, which plays a major role in immune evasion and HIV pathogenesis, exerts several functions during HIV infection (for review, see Cullen, 1998; Fackler et al., 2007; Malim and Emerman, 2008). Nef is responsible for the down-regulation of cell surface CD4 (Aiken et al., 1994) and MHC class I (Le Gall et al., 1998), resulting in the inhibition of antigen presentation to CTL and a decrease of CTL-mediated cell lysis (Collins et al., 1998). Nef expression was also shown to disturb several signalling and trafficking pathways (Roeth and Collins, 2006). Finally, Nef was shown to play a key role in T-cell activation. In most lentiviruses, Nef inhibits T-cell activation; however, this function was lost in the lineage that gave rise to HIV-1 (Schindler et al., 2006), contributing to the increased pathogenesis of HIV-1 compared to HIV-2 or SIV. Indeed, Nef has the ability to slow down the internalisation and recycling of TCR-CD3 complexes, impairing their accumulation at the immunological synapse - site of antigen-presentation and T-cell activation - and therefore reducing T-cell activation (reviewed by Fackler et al., 2007). However, Nef proteins from 'pathogenic' viruses such as HIV-1 do not down-regulate cell surface levels of CD3, resulting in an increased activation sensitivity of T-cells and cell death (Schindler et al.,

2006). Some exceptions to this pattern exist as HIV-2 and SIV_{MAC} Nef induce the decrease of CD3 at the cell surface despite these viruses being pathogenic (Fackler et al., 2007).

The main role of HIV Vif is to block the antiviral activity of human APOBEC3G and APOBEC3F (Malim and Emerman, 2008). APOBEC3G is a cellular cytidine deaminase expressed in cell lines such as H9 and CEM (non-permissive to HIV infection), which inhibits HIV replication using cytidine deaminase-dependent and -independent mechanisms (Bishop et al., 2008; Holmes et al., 2007; Sheehy et al., 2002). HIV Vif prevents the incorporation of APOBEC3G into nascent particles and targets APOBEC3G to the proteasome where it is degraded (Sheehy et al., 2003).

Vpr is packaged into HIV virions due to its interaction with the p6 domain of Gag (Cullen, 1998). Vpr contains a NLS and is involved in the nuclear import of the PIC, particularly in non-dividing cells (Heinzinger et al., 1994). In addition, Vpr has been shown to be cytopathic for cells and to delay or block the G2 to mitosis transition of the cell cycle (Malim and Emerman, 2008). Vpr was found to interact with DCAF1 (DDB1-Cul4A-associated factor), which is a receptor of the Cul4A (cullin 4A)-DDB1 (damaged-DNA-specific binding protein 1) ubiquitin (Ub) ligase (Le Rouzic et al., 2007), and this interaction appeared critical for its ability to arrest the cell cycle in G2 (Dehart and Planelles, 2008).

In HIV-1, Vpr also facilitates infection of macrophages, a function assumed by the Vpx protein in HIV-2 and SIV_{SM}. Indeed, HIV-2 and SIV_{SM} infection of macrophages or monocyte-derived DCs is dependent upon Vpx expression (Goujon et al., 2007; Sharova et al., 2008), suggesting that Vpx helps overcome a (yet unknown) restriction mechanism in macrophages. Interestingly, Vpx interacts with the same cul4A-DDB1-

DCAF1 complex as Vpr (Le Rouzic et al., 2007), possibly targeting that unknown restriction factor to the proteasome for degradation.

1.5 The HIV life cycle

Many aspects of the life cycle of HIV-2 are similar to that of HIV-1. An overview is presented in Figure 1.6.

1.5.1 Receptor binding and entry

HIV-1 enters the target cell through the binding of its surface Envelope glycoprotein (gp120) to the CD4 cell surface marker (Dalgleish et al., 1984). Binding to a co-receptor is also required and CCR5 and CXCR4, two proteins from the family of seven-transmembrane-domain G-protein-coupled receptors, were identified as HIV-1 co-receptors (Deng et al., 1996; Dragic et al., 1996; Feng et al., 1996a). Macrophage-tropic HIV strains use CCR5, whereas T-cell-tropic strains of HIV use CXCR4. This difference in co-receptor usage explains the resistance of certain high risk populations to HIV-1, as it was found that those individuals lack CCR5 (Liu et al., 1996; Samson et al., 1996).

Unlike HIV-1, primary isolates of HIV-2 have been shown to use a wide range of co-receptors, including CCR1, CCR2b, CCR3, CCR4, CCR5, and CXCR4 (McKnight et al., 1998). In addition, HIV-2 differs slightly from HIV-1 in that it can enter the cell in a CD4-independent manner, using CXCR4 as a main receptor (Endres et al., 1996; Reeves et al., 1997).

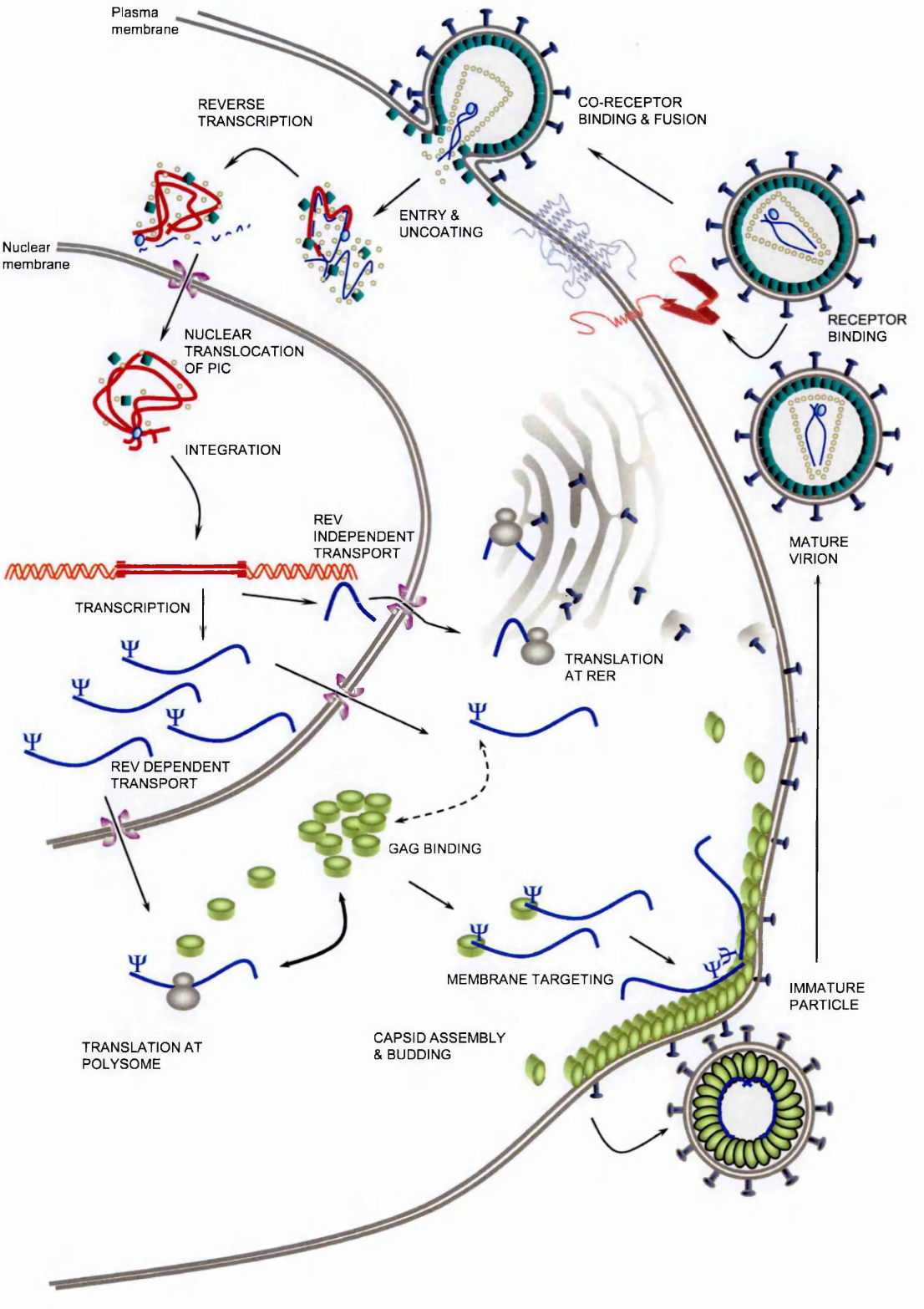


Figure 1.6: The HIV-2 life cycle.
Abbreviations are as defined in the text. Figure courtesy of S.D. Griffin.

HIV entry occurs through the binding of a gp125 trimer to CD4 followed by fusion of the viral envelope to the plasma membrane.

In addition to the virus-cell interaction-mediated entry, HIV is efficiently transmitted from cell to cell, using the virological synapse formed upon contact between an infected and an uninfected cell (Hubner et al., 2009). It is believed that cell-cell HIV transmission is more efficient than virus-cell transmission, probably because it protects the virus from the immune system (Martin and Sattentau, 2009).

1.5.2 Reverse transcription

Following the release of the capsid into the cytoplasm of the cell, the viral genome is stripped of its CA and MA layers in a process known as uncoating. As discussed in section 1.4.2, CypA is thought to play a role in uncoating the viral genome. The remains of the viral core, a nucleoprotein complex composed of the RNA genome associated with RT, IN, Vpr and Vpx, a few MA molecules as well as a small amount of CA and NC, constitutes the pre-integration complex or PIC (Fouchier and Malim, 1999), which protects the viral RNA and proviral DNA from nucleases (Tanchou et al., 1995) and constitutes the site of the reverse transcription. The RT enzyme and $tRNA^{lys3}$ primer are provided by the incoming virion. A schematic of reverse transcription is presented in Figure 1.7.

Reverse transcription takes place in the cytoplasm of the cell and starts with the synthesis of the minus strand (-) strong stop (ss) cDNA from the tRNA primer (Figure 1.7, step 1). Once the RT enzyme reaches the 5′ end of the genome, its RNase H activity triggers the degradation of the RNA in the RNA-DNA duplex, *i.e.* RNA sequences from the 5′ end to the PBS (step 2).

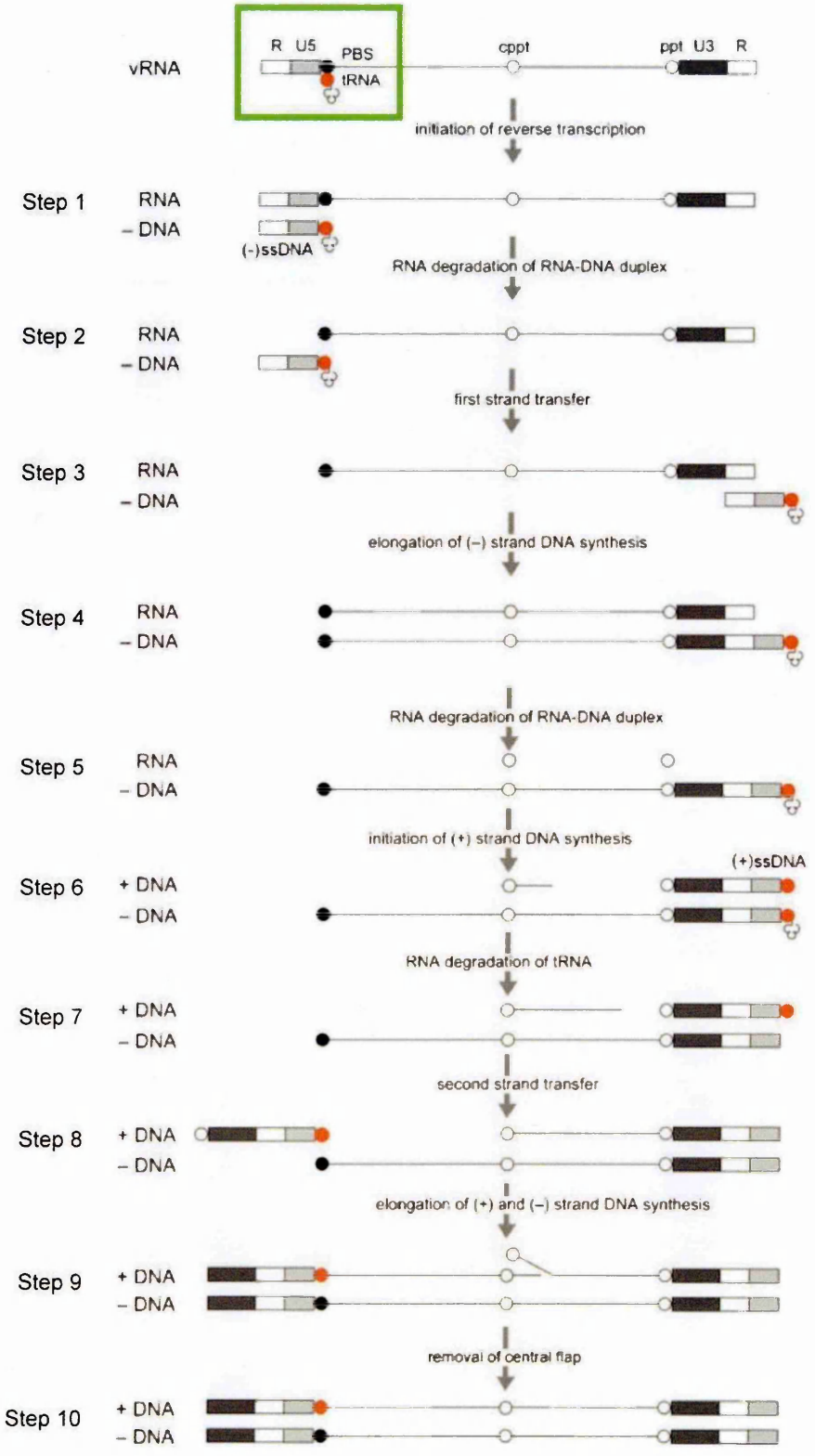


Figure 1.7: Mechanism of reverse transcription.
Adapted from Abbink and Berkhout (2007).

The (-) ssDNA is then transferred to the 3′ end of the genome, where its R region anneals to the complementary R sequence on the viral RNA (step 3). This first strand transfer may be inter- or intra-molecular (van Wamel and Berkhout, 1998), which allows the virus to recombine during reverse transcription (Hu and Temin, 1990b). The (-) ssDNA is elongated from the R region to the PBS (step 4) and the RNA template from the RNA-DNA is degraded (step 5), with the exception of two RNase-resistant regions, the central and 3′ polypurine tracts (cPPT and PPT respectively) which serve as primer for the positive (+) strand DNA synthesis. Following the initiation of (+) strand DNA synthesis from the two polypurine tracts (step 6), the tRNA primer is degraded (step 7) and the 3′ extremity of the (+) strand DNA jumps to the 5′ end of the (-) strand DNA and anneals to the 5′ copy of the PBS (second strand transfer, step 8). Finally, the (+) and (-) strand DNAs synthesis is completed (step 9), creating a central flap on the (+) strand DNA at the site of the cPPT that is subsequently removed (step 10). The proviral DNA now contains a copy of the U3-R-U5 region, forming the LTR, at each extremity of the genome. During or after reverse transcription, the PIC is transported into the nucleus of the cell, through interaction of the NLS of certain PIC components with the nuclear pores complex. Nuclear import of the PIC is thought to be mediated mainly by the Vpr protein for HIV-1 (Mahalingam et al., 1997) and the Vpx protein for HIV-2 during infection of non-dividing cells (Fletcher et al., 1996; Mueller et al., 2004; Pancio et al., 2000).

1.5.3 Integration

Once the PIC has been imported in the nucleus, the proviral DNA can integrate into the genome of the host cell, a step that is necessary for a productive HIV infection.

Integration occurs in three steps (reviewed in Delelis et al., 2008; Vandegraaff and Engelman, 2007). First, an IN tetramer binds to the LTRs on the linear proviral DNA and catalyses the 3′ endonucleotide cleavage of a dinucleotide, known as 3′ processing, generating 3′-hydroxyl groups and 5′ overhangs. The cleaved DNA is then integrated into weakly conserved palindromic sequence on the host chromosome by a DNA strand-transfer reaction catalysed by the integrase. The 3′-hydroxyl groups is used by IN to cut the 5 base pairs (bp) target and the 3′ viral ends are joined to the 5′-phosphate of the cut, creating a DNA recombination intermediate with joined viral 3′ ends and free 5′ ends. Finally, DNA repair of the DNA recombination intermediate yields a 5 bp duplication of the target DNA flanking the integrated provirus.

Integration of HIV-1 and HIV-2 has been shown to occur preferentially in transcriptionally active sites of the host genome, with high gene density and high GC content (MacNeil et al., 2006; Mitchell et al., 2004; Schroder et al., 2002). Interestingly, the rate of integration into heterochromatin, which results in a block of viral replication, was found to be higher in HIV-2 compared to HIV-1 infected individuals (MacNeil et al., 2006), possibly contributing to the lower pathogenicity of HIV-2.

For HIV-1, the preferential integration in transcription units was found to be lost in cells depleted of a cellular protein called lens epithelium-derived growth factor p75 or LEDGF/p75 (reviewed by Engelman and Cherepanov, 2008). LEDGF/p75 is a cellular co-factor of lentiviral integrases and has been shown to stimulate the binding of the integrase to the chromatin (Busschots et al., 2005).

The integrated provirus may remain latent or is transcribed by the cellular machinery as described below. Latency is an important aspect of HIV pathogenesis that is affected by the site of integration and *cis*- and *trans*-acting factors (Han et al., 2007; Lassen et al., 2004; Marcello, 2006). Although reactivation of the HIV genome remains poorly

understood, it is believed that several factors such as T-cell activation and nuclear factor kappa B (NF-κB) are involved (Mok and Lever, 2008).

1.5.4 Synthesis and processing of viral RNAs

Viral mRNA transcripts are synthesised by the cellular machinery, namely the RNA polymerase II. The U3 region of the 5′ LTR constitutes the promoter region and as such binds a series of factors that regulate transcription (Kingsman and Kingsman, 1996). For example, two binding sites for NF-κB are present in a region called the transcriptional enhancer. Binding of the transcription factors TFIID (a complex of TATA-binding protein [TBP] and TBP-associated factor [TAF]) and SP1 also results in a stimulation of HIV transcription. On the other hand, NFAT-1 (nuclear factor of activated T-cells 1) and USF (upstream stimulatory factor) recognise sequences located in the *cis*-acting negative regulatory element and inhibit HIV transcription.

As described in section 1.4.2, the viral protein Tat also plays an important role in stimulating HIV gene expression. To do so, Tat binds to the TAR stem-loop in the R region of the 5′ LTR, recruits the two components of the protein kinase P-TEFb, cyclin T1 and CDK9, thereby promoting the phosphorylation of Pol II CTD and enhancing its processivity. It was shown that recruitment of the P-TEFb kinase to the promoter region was the sole role of the Tat-TAR interaction as artificial recruitment of P-TEFb was both necessary and sufficient to activate LTR-driven gene expression in the absence of Tat and TAR (Bieniasz et al., 1999).

Following transcription, viral mRNAs encoding for Tat, Rev and Nef are fully spliced and exported out of the nucleus by the cellular machinery. The situation is slightly more complex for other viral mRNAs such as Env mRNA and the genomic RNA, which are incompletely spliced but still need to be exported to the cytoplasm for Env, Gag and

Gag-Pol expression as well as genome encapsidation. Intron-containing RNAs remain in the nucleus where they are normally targeted for degradation. However, the Rev protein ensures the nuclear export and expression of incompletely spliced mRNAs, as detailed in section 1.4.2.

1.5.5 Synthesis and processing of viral proteins

Once in the cytoplasm of the cell, viral mRNAs are transported either to the free polyribosomes or to the rough endoplasmic reticulum (RER) to be translated by the cellular machinery (Coffin et al., 1997).

- ***Translation of the Envelope glycoprotein***

Cap-dependent translation of Env starts on free ribosomes and shortly thereafter the leader (L) peptide emerges. This hydrophobic peptide is recognised by the signal recognition particle (SRP), which docks onto the membrane of the RER, and is inserted into the lipid bilayer. The Env protein is cotranslationally translocated across the RER membrane, where two modifications occur. First, the L peptide is proteolytically removed by a cellular signal peptidase and second, the protein is glycosylated. The synthesis of the hydrophobic cytoplasmic tail prevents the release of Env into the lumen of the RER and secretion from the cell, allowing the protein to be correctly folded and assembled into trimers prior to transport to the plasma membrane. The folded protein is able to interact with its receptor, CD4, and the Vpu protein of HIV-1 is known to degrade CD4 to prevent early interaction events (Willey et al., 1992). In HIV-2, gp140 itself and Nef were proposed to limit premature contact between Env and CD4 by targeting CD4 for degradation (Cucchiarini et al., 1995; Lindwasser et al., 2007),

although the role of gp140 on CD4 destabilisation was not confirmed (Bour and Strebel, 1996). Env is trafficked to the cell surface through the Golgi network and the secretory pathway, and is subjected to further post-translational modifications during this process.

- ***Translation of the Gag and Gag-Pol precursors***

The Gag and Gag-Pol polyproteins are translated from the unspliced genomic RNA on free polysomes in the cytosol of the cell. To date, there is still controversy as to which translation initiation mechanism is used to produce these proteins. While some studies have proposed that Gag and Gag-Pol translation is cap-dependent and occur through ribosome scanning of the 5′ UTR, there is evidence suggesting that Gag translation is governed by an internal ribosome entry site (IRES) activity (Yilmaz et al., 2006). In particular, the regulatory elements present in the 5′ UTR of many lentiviruses constitute strong structural barriers for the ribosome to overcome, for which a cap-independent translation mechanism would provide an interesting alternative. In both HIV-1 and HIV-2, IRES have been identified within the *gag* coding sequence (Buck et al., 2001; Herbreteau et al., 2005). In addition, Brasey and co-workers identified an IRES in the HIV-1 5′ UTR that was active during the G2/M phase (Brasey et al., 2003), when the efficiency of cap-dependent translation is reduced. However, two other studies showed that the HIV-1 5′ UTR inhibited translation of a *gag* reporter construct, suggesting that it did not contain an IRES (Buck et al., 2001; Miele et al., 1996). As a result, it has been proposed that HIV-1 uses both cap-dependent and -independent initiation mechanisms, ensuring optimal Gag production in response to changes in the cellular environment and progression of HIV infection (Yilmaz et al., 2006). However, it is not clear whether a similar coordinate use of ribosome scanning and IRES activity exists in HIV-2. Three IRES have been identified in the *gag* coding region, which direct the translation of three

Gag isoforms, p57, p50 and p44 (Herbreteau et al., 2005), and it was proposed that high concentrations of Gag could inhibit the IRES-mediated translation of p57 Gag by binding to the HIV-2 5′ UTR (Ricci et al., 2008). Interestingly, self-regulation of its own translation by the Gag protein has been observed with HIV-1 (Anderson and Lever, 2006). In this case, low concentrations of Gag resulted in a stimulation of translation, mediated by the MA domain of Gag, and high Gag concentrations triggered an inhibition of Gag translation that was dependent on the binding of NC to the packaging signal. These observations, together with the lack of separate 'translation' and 'packaging' pools of HIV-1 and HIV-2 genomic RNA (Dorman and Lever, 2000), suggest that there is a competition between HIV translation and packaging in order to achieve the correct balance between production of the structural protein Gag and assembly of new particles.

Gag is modified cotranslationally at its N-terminus by the addition of a myristate on the first glycine residue, which is critical for the membrane targeting of Matrix. In addition, the matrix domain of Gag is phosphorylated, a modification that is thought to reveal the NLS in MA and to play a role in the PIC nuclear import (Bukrinsky et al., 1993a; von Schwedler et al., 1994).

The Gag-Pol polyprotein is synthesised following a -1 ribosomal frameshifting event, as described in section 1.4.2.

- ***Proteolytic processing***

The Env precursor gp140 is cleaved by a cellular protease into gp125 and gp36. Gag and Gag-Pol polyproteins are cleaved after particle assembly, during the maturation of the virion, by the virally encoded protease, providing the structural proteins MA, CA and NC, p1, p2 and p6, which functions have been detailed in section 1.4.2.

1.5.6 Assembly, release and virion maturation

For assembly to occur, all components of the viral particle must reach the site of assembly, *i.e.* traffic to the plasma membrane. While Env seems to traffic with the secreted cellular proteins through the Golgi network and ER, it is not clear which pathway the genomic RNA and Gag polyprotein utilise to reach the site of particle assembly, and whether these traffic together as a complex or are targeted independently of each other (Klein et al., 2007; Swanson and Malim, 2006).

- ***Intracellular trafficking of HIV Gag***

Due to the presence of a myristate at its N-terminus, the Matrix protein has been suggested to play a role in targeting Gag to the plasma membrane. Binding of MA to the plasma membrane has been shown to require both the myristylation of the N-terminus and downstream basic residues (Ono and Freed, 1999; Paillart and Gottlinger, 1999; Spearman et al., 1994; Zhou et al., 1994) and is determined by the degree of exposure of the myristate moiety, which would be high in the context of the Gag polyprotein and low in the case of the mature MA protein, as described in the myristyl switch model (reviewed in Freed, 1998; Gottlinger, 2001). This model is supported by a nuclear magnetic resonance (NMR) structural analysis of HIV-1 MA that showed that the myristate is sequestered in the context of MA, which is mainly monomeric, and exposed in MA-CA polypeptides, which can trimerise (Tang et al., 2004).

Several studies that used confocal microscopy and FRET (fluorescence resonance energy transfer) have shown that the endosomal pathway was implicated in HIV Gag trafficking to the plasma membrane (Nydegger et al., 2003; Perlman and Resh, 2006). Some reports also proposed that the late endosomal compartment or multivesicular

bodies (MVBs) were the site of HIV-1 assembly in certain cell-types such as macrophages (Nydegger et al., 2003; Ono and Freed, 2004; Pelchen-Matthews et al., 2003; Raposo et al., 2002; Sherer et al., 2003). Recently, it was shown that, under certain circumstances, MVBs could support the productive assembly of an HIV-1 16EK/29/31KE MA mutant in macrophages and T-cells (Joshi et al., 2009). However, whether HIV assembly occurs in MVBs remains a controversial subject, as it was shown that HIV Gag does not interact with cellular markers of MVBs such as CD63 (Gomez and Hope, 2006). Several studies have reported that endocytosis and re-internalisation account for the endosomal localisation of Gag and that plasma membrane is the primary site of assembly (Finzi et al., 2007; Jouvenet et al., 2006). In particular, Jouvenet and co-workers demonstrated that (i) Gag localisation at the plasma membrane preceded its endosomal accumulation in several cell lines including primary macrophages, (ii) microtubule or endosome motility was not required for HIV-1 assembly in macrophages and (iii) particles were not released when assembly was targeted to endosomal compartments. A recent study analysing the assembly of individual HIV-1 virions in live cells supports this model and confirmed that the plasma membrane is the primary site of HIV-1 assembly (Jouvenet et al., 2008). Further studies have also suggested that HIV assembly in macrophages does indeed occur at the plasma membrane. In their report, Deneka et al. proposed that HIV-1 assembles into an intracellular plasma membrane domain, containing the tetraspanins CD81, CD9 and CD53, that is distinct from endosomes (Deneka et al., 2007). Interestingly, these results were corroborated by the analysis of HIV trafficking in monocyte-derived macrophages using a biarsenical labelling system (Gousset et al., 2008). A combination of plasma membrane and apparently internal Gag staining was observed, the latter moving to the site of synapse formation after cell-cell contact with an uninfected T-cell. This suggests

that the tetraspanin-rich intracellular compartment might serve as a reservoir for rapid presentation of virus at cell-cell junctions.

Supporting this model further, the analysis of the cellular localisation of ESCRT components in infected and uninfected T-cells and macrophages revealed that only 3 % of total ESCRT staining localised with the MVBs (Welsch et al., 2006). Mostly, components of the ESCRT machinery were found on tubular-vesicular endosomal membranes. The ESCRT labelling on endosomes was twice as concentrated in macrophages compared to T-cells, while labelling was twice more abundant at the plasma membrane of T-cells, possibly explaining early suggestions that HIV-1 can assemble in late endosomes in macrophages.

- ***Host factors involved in HIV Gag trafficking***

Several host factors have been linked with Gag intracellular trafficking. Cholesterol was shown to be important for Gag mobility in cells (Gomez and Hope, 2006) and Phosphatidylinositol-(4,5)-biphosphate [$PI(4,5)P_2$], a member of the phosphoinositide family of lipids which participates in directing cellular proteins to the plasma membrane (McLaughlin and Murray, 2005), was also found to play an important role in directing Gag to the plasma membrane (Ono et al., 2004). Interestingly, structural studies have demonstrated that $PI(4,5)P_2$ binds to the myristylated matrix domain of HIV-1 Gag, inducing a conformational change and triggering the exposure of the myristate moiety (Saad et al., 2006). This event, called the myristyl switch, has been shown to be critical for the anchoring of MA into lipid bilayers, as discussed above. The role of $PI(4,5)P_2$ in providing an anchor for Gag in the plasma membrane and triggering the myristyl switch helps provide an explanation for the observation that depletion of $PI(4,5)P_2$ induces the relocalisation of HIV-1 assembly to MVBs (Ono et al., 2004).

It has also been suggested that the cytoskeleton is used and HIV-1 Gag was shown to interact with the suppressor of cytokine signalling 1 (Ryo et al., 2008), resulting in an increased level of microtubule-associated Gag (Nishi et al., 2009). Furthermore, depletion of the kinesin family member 4 (KIF4), a microtubule-stimulated ATPase, resulted in an accumulation of Gag in perinuclear clusters, an increased Gag degradation and an inhibition of VLP production (Martinez et al., 2008).

HIV-1 MA has been shown to interact with the δ subunit of the clathrin adaptor protein (AP)-3 complex (Dong et al., 2005), which is involved in the endocytic pathway. Depletion of AP-3 decreases particle release and alters Gag trafficking pattern. The ESCRT-I component Tsg101, which interacts with Gag via the late domain p6, has also been found to control HIV-1 Gag trafficking to the plasma membrane (Goff et al., 2003), with both the p6- and Ub-binding motifs of Tsg101 being required for this function.

The ADP ribosylation factor (Arf) and Golgi-localised γ-ear containing Arf-binding (GGA) proteins were also shown to be involved in Gag trafficking and particle release (Joshi et al., 2008). GGA overexpression disrupted Arf activity, resulting in the inhibition of particle production due to the decreased Gag-membrane binding and the sequestering of Gag in intracellular compartments. This altered Gag trafficking was associated with a reduction in retrovirus particle production. These data suggest that GGA acts as a modulator of HIV-1 release and Arf as a critical cellular cofactor in retroviral Gag trafficking to the plasma membrane.

Finally, the ATP-binding cassette protein from the E subfamily, ABCE1, was also implicated in particle assembly but its exact role remains unclear (Zimmerman et al., 2002).

- ***Trafficking of the genomic RNA to the plasma membrane***

Several studies have tried to determine whether the RNA and Gag traffic independently or together to the plasma membrane. Using confocal microscopy and FRET, Poole and co-workers identified that HIV-1 Gag and genomic RNA interacted at a perinuclear localisation, suggesting that pericentriolar region might be the site at which HIV-1 Gag selectively associates with the genomic RNA (Poole et al., 2005). However, these results were obtained following overexpression and have not been observed by others. Nonetheless, the pericentriolar region has been shown to be the site of MPMV Gag accumulation and immature capsid assembly (Sfakianos et al., 2003), and depletion of the heterogeneous nuclear ribonucleoprotein A2 (hnRNP A2) trafficking factor, which has been implicated in HIV-1 RNA trafficking due to homology between the nucleotide sequence of the MHR in Gag and the hnRNP A2 response element (Mouland et al., 2001), resulted in a similar pericentriolar accumulation of HIV-1 RNA (Levesque et al., 2006). Finally, the dsRNA-binding protein Staufen, which has been proposed to play a role in HIV-1 RNA encapsidation (Mouland et al., 2000), may take part in genomic RNA trafficking since human Staufen is involved in mRNA transport and contains a tubulin-binding domain (Wickham et al., 1999).

To date, there is no evidence to support the idea that HIV genomic RNA traffics to the plasma membrane independently of Gag (reviewed in Swanson and Malim, 2006). Interestingly, it has been proposed that for MLV, pre-budding complexes containing Env, Gag and the genomic RNA were formed on endosomes and subsequently transported to the plasma membrane (Basyuk et al., 2003). One could envisage a similar pattern in HIV involving Gag, Env and the genomic RNA trafficking via the endocytic pathway (Murakami, 2008).

- ***Genome encapsidation and dimerisation***

RNA packaging or encapsidation is the specific capture of the genomic RNA by the structural protein Gag amongst the vast excess of cellular mRNAs and spliced viral RNA species. To achieve this specificity, the NC domain of Gag recognises a highly structural motif in the 5′ UTR of the unspliced RNA, termed the packaging signal or Psi (Ψ), and targets RNA incorporation into nascent virions. The detailed process of genome packaging in lentiviruses will be described in section 1.6.

A unique feature of retroviruses is that each particle contains two copies of the RNA genome. Early studies of type C retroviruses revealed that the two copies of the genome were non-covalently linked at their 5′ extremity, in a region termed the dimer linkage site or DLS (Bender et al., 1978; Bender and Davidson, 1976; Kung et al., 1976). Dimerisation is believed to occur in stages, with the first interaction occurring through Watson-Crick and non canonical bonds between partial or complete complementary palindromic sequences (for review see Paillart et al., 2004b). The process of genome dimerisation will be the subject of a more detailed introduction in section 1.7.

To date, it remains unknown whether the dimerisation of the genomic RNA occurs prior to or after particle release. In either case, two copies of the RNA have to be selectively encapsidated by the Gag polyprotein.

- ***Recruitment of the retroviral Envelope***

As mentioned previously, Env traffics to the plasma membrane via the Golgi/ER and is then recruited to the site of viral budding by a mechanism that remains poorly understood. A model has been proposed in which Gag indirectly interacts with Env via a cellular intermediate, tail-interacting protein of 47 kDa or TIP47 (Lopez-Verges et al.,

2006). TIP47 has been proposed to act as a bridge between HIV-1 Env and MA and disruption of the MA-TIP47 interaction or depletion of TIP47 was reported to abrogate Env incorporation and decrease viral infectivity (Lopez-Verges et al., 2006).

Analysis of the envelope-pseudotyping phenomenon, by which a virus can incorporate an exogenous Env, also suggests that Env recruitment is not a random mechanism, as co-expression of HIV-1 Gag and MLV Env results in a patchy Env staining as opposed to a diffuse staining when Env is expressed alone (Jorgenson et al., 2009). It was hypothesised that viral budding creates a unique lipid environment that would recruit the mature, *i.e.* cleaved, Env protein.

- ***Particle assembly and release***

Once all the elements have reached the correct site of assembly, retroviral particles can assemble. The composition of a typical retroviral particle is approximately as follow: Gag accounts for 47 %, lipids for 30 % and the genome for 2.2 % of the particle mass (Krausslich, 2008). The rest of the particle mass is accounted for by the viral (e.g. Env, IN, RT, Vpr) and cellular proteins (e.g. CypA, actin) that are incorporated in the virion.

An important aspect of particle assembly is the multimerisation of the structural protein Gag, a necessary step considering that approximately 2400 Gag molecules constitute the lattice of a 66 nm retroviral particle (Krausslich, 2008). The Gag domains involved in multimerisation have been described in section 1.4.2. *In vitro* studies showed that CA-NC proteins assemble into VLPs more efficiently than CA alone (Campbell and Vogt, 1995). Furthermore, CA-NC assembles into VLPs *in vitro* in the absence of any other viral elements, even though the presence of nucleic acid strongly enhanced capsid formation and VLP assembly (Campbell and Rein, 1999; Campbell and Vogt, 1995; Ganser et al., 1999). RNAse treatment was found to disrupt Gag multimerisation *in vitro*

(Burniston et al., 1999) and MLV mutants bearing deletion in the packaging signal were shown to encapsidate cellular RNA (Muriaux et al., 2001) providing further evidence that RNA probably plays a role in retroviral assembly. In addition, NC was shown to be required for Gag multimerisation (Burniston et al., 1999; Franke et al., 1994a) and it appeared that the RNA binding domain of NC was necessary for this activity (Cimarelli et al., 2000). Altogether, these studies support the hypothesis that the NC domain of the Gag precursor binds to and recruits either genomic or cellular RNA, which has been proposed to act as a scaffold for the assembly of the capsid (Campbell and Rein, 1999; Campbell and Vogt, 1997; Ganser et al., 1999; Hogue et al., 2009; Muriaux et al., 2001; Muriaux et al., 2002; Rulli et al., 2007).

HIV particle formation is believed to occur in a single assembly event as the hexameric lattice that forms the capsid is in one piece, even though it is incomplete (Briggs, 2008). Despite great advances in the past few years to decipher and understand the morphogenesis of HIV particles, many aspects remain to be solved.

In recent years, great efforts have been developed to understand the process of HIV release and many cellular cofactors have been characterised. However, many aspects still remain unclear. It has now been clearly established that HIV 'hijacks' the components of the ESCRT machinery to assemble at the plasma membrane and bud from the cell (reviewed in Fujii et al., 2007; Martin-Serrano, 2007; Morita and Sundquist, 2004).

As mentioned in section 1.4.2, the p6 domain of Gag interacts with the ESCRT-I component Tsg101 and the protein AIP1/Alix (Garrus et al., 2001; Martin-Serrano et al., 2001; Strack et al., 2003). ESCRT-I is a component of the machinery that promotes the vesiculation of the endosomal membrane during cellular budding, a process that resembles enveloped-virus budding, by sorting ubiquitinated transmembrane proteins

(cargo) into MVB vesicles. This ATP-dependent process is also dependent on ESCRT-II and ESCRT-II complexes. The p6-Alix interaction is thought to promote particle release by connecting Gag with components of the ESCRT-III complex, such as chromatin modifying protein 4 (CHMP4; von Schwedler et al., 2003b). This hypothesis is supported by the fact that rescue of HIV-1 late domain mutants by AIP1/Alix depends upon its ability to bind CHMP4 (Fisher et al., 2007; Usami et al., 2007).

Particle release is mediated by the ESCRT machinery and enhanced by the Vpu protein in HIV-1 and Env in HIV-2. Recent studies have indeed demonstrated that these two proteins function to overcome factors such as tetherin (Neil et al., 2008) and the calcium-modulating cyclophilin ligand (Varthakavi et al., 2008) that restrict the release of HIV.

- ***Particle maturation***

Immature particles are released from the cell and subsequently undergo a protease-dependent maturation event. The viral protease is self-cleaved from the Gag-Pol precursor and proteolytically processes the Gag precursor and the remaining of Gag-Pol to release the mature forms of MA, CA, NC, RT and IN. These cleavages trigger a conformational change in the virions, the capsid forming a conical core underneath the MA-layered lipid membrane. Based on *in vitro*-assembled capsids, a model has been proposed for the mature HIV particle assembly, named the 'fullerene' model, in which viral cores are organised on conical hexagonal lattices composed of hexameric CA rings with five pentamers at the narrow end and seven at the wide end, with a cone angle of 19.2 degrees (Ganser et al., 1999; Li et al., 2000). A similar hexagonal unit has been observed by cryo-electron microscopy (cryo-EM) and thin-section electron microscopy for authentic mature HIV-1 cores (Briggs et al., 2003; Welker et al., 2000), suggesting

that the hexameric unit seen in the *in vitro*-assembled CA tubes is relevant to mature HIV particles.

Inside the core, the dimeric RNA genome is coated with NC. It has been demonstrated that the RNA dimer in the mature particle is more stable than RNA dimers in immature particles. Studies of protease-defective viruses have shown that the electrophoretic mobility, density-sedimentation profile and thermal stability of these dimers were different from that of wild type dimers (Fu et al., 1994; Fu and Rein, 1993; Oertle and Spahr, 1990; Stewart et al., 1990). The NC protein is thought to promote this maturation of the dimer, by triggering the transition from a loose dimer to an more stable dimer (reviewed in Paillart et al., 2004b), as described in section 1.7.2.

1.6 Genome encapsidation in retroviruses

1.6.1 *Cis*-acting RNA packaging signal

The genome of retroviruses represents less than 1 % of total RNA present in an infected cell (Coffin et al., 1997). Hence, retroviruses have developed highly specific mechanisms to selectively package their genomic RNA over the vast excess of cellular RNA on one hand, and over spliced viral RNA species on the other hand (reviewed in D'Souza and Summers, 2005; Lever, 2007).

Studies of simple retroviruses have revealed that the *cis*-acting RNA sequences required for packaging were located in the 5′ UTR of the genome (Katz et al., 1986; Mann et al., 1983; Sorge et al., 1983; Watanabe and Temin, 1982) and it was supposed that the RNA sequence required for HIV genome encapsidation would also be located in this region. The location of the packaging determinant Psi (Ψ) 3′ of the major splice donor would ensure that only unspliced genomic RNA was incorporated into nascent virions.

Mutagenesis analysis of the 5′ leader RNA of HIV-1 revealed that a 19 nt region, located downstream of the major splice donor, was necessary for efficient RNA packaging (Lever et al., 1989). This was confirmed by further studies which also mapped the packaging signal to the region between the SD and the *gag* AUG start codon (Aldovini and Young, 1990; Clavel and Orenstein, 1990). Interestingly, as well as decreasing overall RNA packaging, Psi mutations also reduced the specificity of encapsidation and spliced viral RNAs were packaged in the mutant virions (Luban and Goff, 1994). These Psi deletions also resulted in a reduction of viral infectivity (Aldovini and Young, 1990; Clavel and Orenstein, 1990; Lever et al., 1989), and in some cases, aberrant particle morphology was observed (Aldovini and Young, 1990; Clavel and Orenstein, 1990). However, none of the deletions introduced resulted in a complete abolition of RNA packaging, suggesting that the packaging determinant may be multipartite. Supporting this hypothesis, the first 40 nt of the *gag* gene were found to be important for HIV-1 encapsidation (Luban and Goff, 1994). The introduction of 5′ *gag* sequences in an HIV-1 vector was also found to enhance gene transfer by thirtyfold and increase RNA encapsidation by fourfold (Parolin et al., 1994). However, it is interesting to note that translation of Gag does not enhance HIV-1 encapsidation, as introduction of a stop codon in CA had no effect on the packaging efficiency of an HIV-1 vector (McBride et al., 1997). Deletions of sequences 5′ of the SD, which overlap with the RNA sequence shown to be involved in genome dimerisation (Darlix et al., 1990; Laughrea and Jette, 1994; Skripkin et al., 1994), have been shown to affect HIV-1 encapsidation (Clever and Parslow, 1997; Kim et al., 1994; McBride and Panganiban, 1996, 1997). Sequences in the TAR, Poly(A) and PBS stem-loops were also found to affect HIV-1 RNA encapsidation to some extent (Helga-Maria et al., 1999; McBride et al., 1997). Overall, it appears that the major HIV-1 *cis*-acting packaging determinant

lies approximately 120 nt upstream of the *gag* AUG and up to 40 nt into the *gag* coding sequence, with some elements at the 5′ extremity of the genome also influencing the packaging process.

Despite having only one pool of RNA for packaging and translation (Dorman and Lever, 2000), HIV-1 can efficiently package RNA vectors that do not produce Gag, using a *trans*-packaging mechanism (McBride et al., 1997). In fact, HIV-1 was shown to use both a *trans*- and a *cis*-packaging mechanism (Kaye and Lever, 1998), even though it was proposed to preferentially package its genome in *cis*, encapsidation being enhanced when the RNAs encode the Gag protein (Poon et al., 2002).

1.6.2 Secondary structure of the HIV-1 packaging signal

The structure of the HIV-1 5′ UTR has been determined by biochemical probing, mass spectrometry, free-energy minimisation and phylogenetic analysis (Baudin et al., 1993; Clever et al., 1995; Harrison and Lever, 1992; Wilkinson et al., 2008; Yu et al., 2008).

The first extensive study of the HIV-1 leader secondary structure came from Harrison and Lever (Harrison and Lever, 1992). The HIV-1 5′ UTR was shown to be highly structured and to form several stem-loop structures between the PBS and the start of the *gag* coding sequence. In particular, this study showed that Psi folds into one of these stem-loops and that the 19 nt deletion described by Lever et al. (1989) disrupted this stem-loop as well as surrounding structures. Baudin et al. also analysed the structure of the HIV-1 leader using biochemical modifications and found that the SD and Psi folded into structures similar to those described by Harrison and Lever (Baudin et al., 1993). Further biochemical probing by Clever and co-workers confirmed the formation of four stem-loops and reported that the sequences creating the stems were particularly conserved (Clever et al., 1995). More recently, mass spectroscopy and SHAPE

(selective 2′-hydroxyl acylation analysed by primer extension) were applied to the HIV-1 Psi RNA (Wilkinson et al., 2008; Yu et al., 2008). These methods confirmed the existence of the four stem-loop structures previously identified and revealed that the compact cloverleaf morphology exhibited by the Psi region is stabilised by long-range tertiary interactions.

Taken together, these independent studies identified a first stem-loop upstream of the SD named stem-loop 1 (SL-1) that comprises the dimerisation initiation signal (DIS), a second stem-loop containing the SD (SL-2), a Ψ-containing stem-loop (SL-3) and a stem-loop immediately downstream of the *gag* AUG (SL-4).

The HIV-1 RNA has also been proposed to contain several long-distance interactions. A compactly folded, fast migrating conformer was identified and addition of Mg^{2+} or NC - two factors that promote RNA dimerisation - was found to destabilise it (Berkhout and van Wamel, 2000). This conformer contains a long-range base-pairing between the Poly(A) and the DIS and was termed the long-distance interaction or LDI structure (Huthoff and Berkhout, 2001). Two mutually exclusive alternate conformations were subsequently proposed for the HIV-1 5′ leader RNA, the LDI and branched multiple hairpin (BMH; Figure 1.8A) structures, which differ in their ability to form RNA dimers, consistent with the effect of Mg^{2+} and NC on the LDI stability (Huthoff and Berkhout, 2001). A long-range interaction between the *gag* AUG and a sequence upstream of the PBS, referred to as the U5-AUG duplex and first identified in the BMH model, was subsequently confirmed by biochemical probing (Abbink and Berkhout, 2003) and is supported by NMR studies (Spriggs et al., 2008). Interestingly, only the full-length genomic RNA can form the dimer-competent BMH structure as the U5-AUG duplex only exist in the unspliced RNA, and it has been suggested that alternate RNA conformations may help regulate processes such as encapsidation (Abbink and

Berkhout, 2003). Supporting this hypothesis, analysis of the LDI-BMH equilibrium of a large set of HIV-1 mutants showed that a LDI-BMH riboswitch regulated encapsidation and dimerisation of HIV-1 RNA (Ooms et al., 2004).

Another intramolecular interaction was identified between the poly(A) loop and nucleotides in the *matrix* coding region (Paillart et al., 2002). Sequences involved in this long-range pseudoknot are conserved in HIV-2 and SIV, suggesting that it is functionally important. It is worth noting that similarly to the U5-AUG duplex, the poly(A)-matrix interaction only exists in the BMH conformer, since the LDI structure does not sustain the poly(A) loop.

Phylogenetic analysis of twenty divergent HIV-1 strains, combined with biochemical probing, confirmed both the poly(A)-matrix and the U5-AUG long-range interactions (Damgaard et al., 2004). Interestingly, the data obtained were consistent with the BMH structure but did not support the LDI model. Furthermore, structural probing of HIV-1 RNA in infected cells and virions confirmed the existence of the dimer-competent BMH structure but not that of the LDI, although the LDI conformer may exist as a minor transient RNA species (Paillart et al., 2004a).

Finally, high throughput SHAPE analysis of the HIV-1 RNA strongly supported the idea that the HIV-1 genomic RNA formed a single predominant structure that included the U5-AUG and poly(A)-matrix interaction and fitted the BMH model (Wilkinson et al., 2008).

The structure of Psi was shown to be important for encapsidation as HIV-1 packaging was severely impaired by disruption of SL1 and SL3 but could be rescued by compensatory mutations that restored the base-pairing (McBride and Panganiban, 1996).

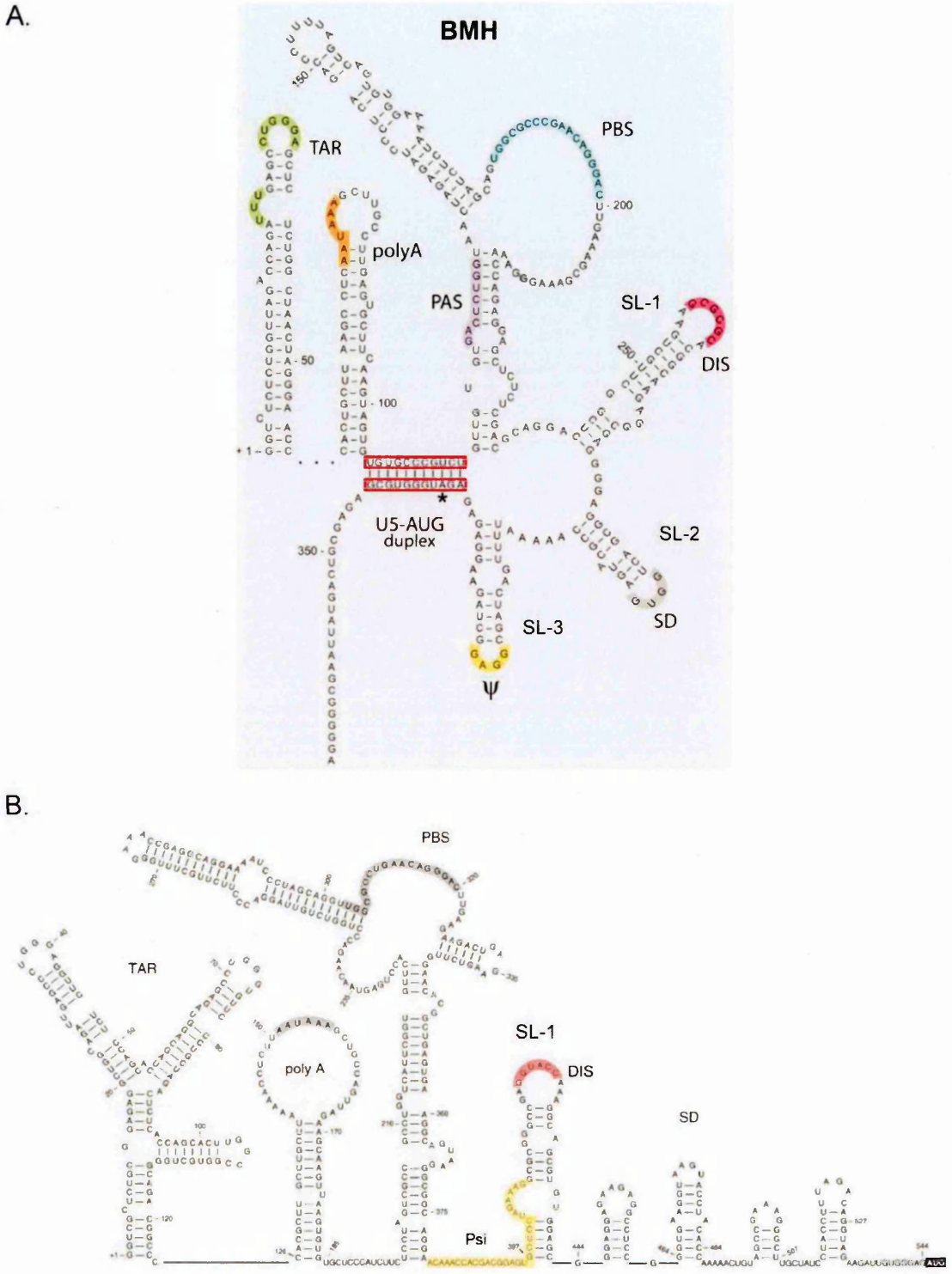


Figure 1.8: Secondary structure of the HIV-1 and HIV-2 5′ RNA leader.

(A) Branched multiple hairpin (BMH) structure of HIV-1_{LAI} 5′ leader RNA from Abbink and Berkhout (2003). The U5-AUG long-range interaction is indicated by the red box. (B) Secondary structure of the HIV-2_{ROD} 5′ leader RNA, adapted from Berkhout (1996).

Residue +1 corresponds to the start of the RNA genome. The proposed dimerisation and encapsidation signals are highlighted in red and yellow, respectively. TAR, *trans*-activation responsive; PolyA, polyadenylation signal; PBS, primer binding site; PAS, primer activation signal; DIS, dimerisation initiation site; SD, splice donor; Ψ/Psi, packaging signal; SL-1/2/3, stem-loop 1/2/3.

Moreover, the base-pairing of the 46 nt located between the SD and *gag* AUG start codon into a stem-loop structure was shown to be crucial for efficient encapsidation (Hayashi et al., 1992), confirming the importance of the Psi structure in HIV-1 RNA packaging. However, both SL-2 and SL4 appeared dispensable for packaging (McBride and Panganiban, 1996).

The structural context of HIV-1 Psi also appears critical for encapsidation. Permutation of SL1 and SL3 showed that the individual structures that compose the HIV-1 Psi function in a position-dependent manner (McBride and Panganiban, 1997). This suggests that the formation of a high-order RNA structure functions to package the full-length genomic RNA and exclude subgenomic RNA species.

1.6.3 *Trans*-acting factors involved in RNA packaging

The role of NC, and in particular of the two CCHC motifs, in retroviral RNA packaging was first identified when mutations in NC resulted in a packaging defect (Aldovini and Young, 1990; Berkowitz and Goff, 1994; Dorfman et al., 1993; Gorelick et al., 1988; Gorelick et al., 1990; Kaye and Lever, 1999; Meric and Goff, 1989; Meric et al., 1988; Schwartz et al., 1997). Even though the zinc fingers were shown to be the main requirement for genome encapsidation, flanking basic residues were also found to play a role in RNA packaging and binding *in vitro* (Dannull et al., 1994; Poon et al., 1996; Schmalzbauer et al., 1996).

NC was shown to confer the specificity of encapsidation. Chimeric HIV-1 virions containing the MLV NC domain preferentially encapsidated MLV RNA and vice-versa (Berkowitz et al., 1995; Zhang and Barklis, 1995). MLV-HIV-1(NC) chimera selectively packaged unspliced HIV-1 RNA over spliced HIV-1 RNA, while the HIV-1-MLV(NC) chimera packaged significantly more non-viral RNAs than wild type HIV-1,

providing further evidence that NC recognises specifically the genomic RNA during encapsidation. Experiments using protease-deficient vectors demonstrated that the Gag precursor was able to package HIV-1 RNA irrespective of whether it was processed or not (Kaye and Lever, 1996). In addition, HIV-1 Gag (p55) and NC (p15) fusion proteins bind to a Ψ RNA with similar affinities *in vitro*, suggesting that the NC domain of Gag contains all the necessary determinants for efficient Ψ RNA binding (Clever et al., 1995; Dannull et al., 1994).

A number of *in vitro* assays have been used to determine the binding affinity of NC for the Ψ RNA, including electrophoretic mobility shift assay (EMSA), filter binding assay, footprinting and NMR (Amarasinghe et al., 2000b; Berkowitz et al., 1993; Clever et al., 1995; D'Souza and Summers, 2004; De Guzman et al., 1998; Shubsda et al., 2002).

HIV-1 NC does not bind to SL-1 with high affinity (Lawrence et al., 2003) but is able to induce a conformational change during the RNA dimer maturation (Feng et al., 1996b), as described in section 1.7.2. Contrary to what has been suggested by mutational analyses, SL-2 appears to play some role in HIV-1 packaging as it is able to bind NC with high affinity (Amarasinghe et al., 2000a). SL-3, the main HIV-1 packaging determinant, also binds NC tightly, as demonstrated by NMR studies (De Guzman et al., 1998; Zeffman et al., 2000) and footprinting experiments (Damgaard et al., 1998). Interestingly, footprinting performed with both HIV-1 Gag and NC revealed differences in the binding pattern of the two proteins and even though they both bound to SL-3, Gag displayed a much stronger affinity (Damgaard et al., 1998), highlighting the importance of using full-length Gag protein in the *in vitro* binding assays. SL-4 displays only a weak affinity for the NC protein when analysed individually (Amarasinghe et al., 2001). Importantly, it was noted that even though individual structures could bind recombinant

Gag and NC proteins, the entire Ψ region showed a far greater affinity (Berkowitz and Goff, 1994; Clever et al., 1995).

In contrast to HIV-1, few studies have assessed the binding of HIV-2 NC to the Ψ region. HIV-2 NC has been shown to bind to the HIV-2 SL-3 (Matsui et al., 2009) and HIV-1 Gag and NC were reported to bind regions of the HIV-2 5′ leader RNA both upstream and downstream of the major SD (Damgaard et al., 1998).

1.6.4 Genome encapsidation in HIV-2

- ***Cis-acting RNA packaging signal***

Genome encapsidation in HIV-2 is less well characterised than in HIV-1. Initial reports proposed that the HIV-2 *cis*-acting packaging signal was located between the major SD and the *gag* AUG start codon (Arya et al., 1998; Garzino-Demo et al., 1995; Poeschla et al., 1998). However, the deletions introduced were often large and did not allow a precise mapping of the Psi region, and mutations upstream of the SD were not fully characterised (Arya et al., 1998; Poeschla et al., 1998). Moreover, one study used an attenuated HIV-2 strain, HIV-2_{ST}, which may differ in its packaging abilities compared to infectious HIV-2 isolates (Garzino-Demo et al., 1995). Subsequent mutational analyses showed that sequences upstream of the SD were required for efficient encapsidation of HIV-2 genomic RNA, while sequences downstream of the SD were less important (Griffin et al., 2001; Kaye and Lever, 1999; McCann and Lever, 1997). Furthermore, a study of HIV-1 and HIV-2 cross-packaging abilities revealed that an HIV-1 helper virus could package HIV-2 vector in *trans*, but that unspliced and spliced HIV-2 RNA were encapsidated in proportion to their cytoplasmic concentrations, suggesting that HIV-2 Psi is present on all RNA species (Kaye and Lever, 1998).

Mapping of HIV-1 NC and Gag binding sites on the HIV-2 5′ leader RNA showed binding on either side of the SD (Damgaard et al., 1998), confirming the results obtained by Kaye and Lever.

The HIV-2 core packaging determinant was subsequently mapped to nucleotides 380-408 of the HIV-2_{ROD} RNA genome (Figure 1.8B, yellow highlight) and deletion of these residues decreased packaging efficiency to approximately 10 % relative to wild type HIV-2 (Griffin et al., 2001). Biochemical and free-energy minimisation analyses have determined that HIV-2 Psi is partly unstructured and partly folds into the stem of a stem-loop termed SL-1 (Baig et al., 2007; Damgaard et al., 1998; Dirac et al., 2001; Lanchy et al., 2004), which encompasses a putative DIS (Dirac et al., 2001). Deletion of nucleotides 392-434 also resulted in a packaging defect (Griffin et al., 2001; McCann and Lever, 1997), suggesting that the structure of SL-1 might play a role in HIV-2 packaging. A recent study of the HIV-2 Psi region confirmed the relevance of the SL-1 structure in HIV-2 RNA packaging (Lanchy and Lodmell, 2007) as long-term replication of Psi mutants revealed that the formation of an extended SL-1 was necessary for viral replication and positively affected RNA encapsidation.

- ***Co-translational packaging of HIV-2 genomes***

The location of HIV-2 Psi upstream of the splice donor raises the question of the specific encapsidation of unspliced genomic RNA over spliced RNAs. As seen in the cross-packaging analysis, HIV-1 Gag preferentially encapsidates HIV-1 unspliced RNA but is unable to discriminate between HIV-2 spliced and unspliced RNAs (Kaye and Lever, 1998). Interestingly, the HIV-2 helper did not cross-package the HIV-1 vector, indicating nonreciprocal packaging and suggesting that the two viruses may use different packaging mechanisms. Introduction of HIV-1 NC into HIV-2 Gag was

sufficient to allow *trans*-packaging of an HIV-1-based vector by the HIV-2 helper virus, demonstrating that NC is important for the specificity of packaging in HIV-2. It is interesting to note that addition of the HIV-1 p2 domain into the chimera enhanced the packaging efficiency, suggesting a role for the p2 domain in encapsidation (Kaye and Lever, 1998). This implication of p2 in RNA encapsidation has been confirmed by a report that the packaging specificity of an HIV-1 DIS mutant incorporating spliced RNA could be rescued by a compensatory mutation in p2 (Russell et al., 2003b).

Further mutational analyses showed that efficient packaging of an HIV-2-based vector only occurred when the genomic RNA packaged expressed the full-length Gag protein, including a functional NC domain with two intact zinc fingers (Kaye and Lever, 1999). Truncations of the N-terminus of Gag, introduction of a stop codon in CA and point mutations of the two zinc fingers in NC all resulted in the inefficient packaging of the vector RNA, and it was proposed that HIV-2 uses a novel type of encapsidation initiation mechanism, in which HIV-2 unspliced RNA is captured in *cis* by the Gag protein, hence providing the specificity for full-length genomic RNA (Kaye and Lever, 1999).

Contrary to what has been observed with HIV-1, HIV-2 cannot efficiently encapsidate its genome in *trans* (Griffin et al., 2001; Kaye and Lever, 1998). However, a helper virus that does not harbour a functional Psi sequence can *trans*-package a vector containing a stop codon in Gag, indicating that the HIV-2-based vector is able to compete for the Gag protein in *trans* (Griffin et al., 2001). Encapsidation efficiency of wild type HIV-2 was assessed in the presence of Psi mutants that could or could not produce functional Gag protein. The packaging efficiency increased only when the wild type construct was co-transfected with a Psi mutant that could synthesise Gag, suggesting that the Gag protein is a limiting factor for HIV-2 encapsidation (Griffin et

al., 2001). Therefore, co-translational packaging and competition for the limiting Gag protein confer specificity of HIV-2 encapsidation.

- ***Similarities and differences with SIV encapsidation***

Sequences important for SIV_{MAC} packaging have also been mapped to the region upstream of the major SD, as sequences on either side of SIV_{MAC239} SD were shown to be important for specific encapsidation (Patel et al., 2003; Whitney and Wainberg, 2006). One study proposed that SIV Psi was located between the SD and the *gag* start codon; however, mutations 3′ of the SD were not analysed independently of those 5′ of the SD, rendering the interpretation of the data difficult (Patel et al., 2003). Other reports have proposed that the core packaging signal lie 5′ of the SD (Guan et al., 2000; Guan et al., 2001b; Strappe et al., 2003) and cross-packaging experiments using HIV-1 helper viruses showed that deletion 5′ of the SD reduced the encapsidation of SIV-based vectors (Rizvi and Panganiban, 1993). Deletions of nucleotides 371-397 and 398-418 of the SIV_{MAC239} genome reduced packaging efficiency to 21 and 44 %, respectively, relative to wild type (Guan et al., 2000; Guan et al., 2001b). Furthermore, deletion of nucleotides 397-429 of the SIV_{MAC239} genome resulted in a packaging efficiency less than 10 % and a loss of packaging specificity, although neighbouring sequences were also shown to affect SIV packaging (Strappe et al., 2003).

Interestingly, the SIV core packaging signal is predicted to fold into a secondary structure similar to that of HIV-2 Psi (Guan et al., 2001b; Strappe et al., 2003) and sequence analysis revealed almost complete conservation between HIV-2 and SIV Psi (HIV sequence compendium 2008). The location of SIV_{MAC} Psi upstream of the major SD would suggest a packaging mechanism similar to that of HIV-2. However, a nonreciprocal packaging mechanism was found to exist between SIV_{MAC} and HIV-2,

with SIV_{MAC} being able to deliver an HIV-2-based vector whereas HIV-2 could not efficiently encapsidate an SIV-based vector in *trans* (Strappe et al., 2005). Interestingly, HIV-1 could efficiently deliver both SIV_{MAC}- and HIV-2-based vectors.

Finally, it is worth noting that not all SIVs are similar regarding their packaging mechanisms; for example, SIV_{AGM} has been shown to encapsidate its genome predominantly in *trans* (Fu et al., 2007).

1.7 Genome dimerisation in retroviruses

1.7.1 The dimer linkage site

Retroviruses encapsidate two copies of their positive sense RNA genome, non-covalently linked at their 5′ end through the dimer linkage site (reviewed in Greatorex and Lever, 1998; Paillart et al., 2004b).

Deletion analysis of the HIV-1_{MAL} 5′ leader RNA revealed that a 100 nt sequence encompassing the Psi region was necessary and sufficient to promote RNA dimerisation *in vitro* in the absence of protein (Darlix et al., 1990). Similarly, deletion mutagenesis of the MLV 5′ UTR showed that the *cis*-acting sequence required for RNA dimerisation *in vitro* overlapped with the packaging signal (Prats et al., 1990). In both studies, the addition of NC enhanced dimer formation and the protein was found tightly associated with the RNA dimer. Introduction of the MLV Psi sequence into non-viral mRNAs resulted in the encapsidation of these RNAs as dimers, supporting the idea that the DLS overlaps with Psi (Hibbert et al., 2004). Further analysis confirmed that nucleotides 311-415 of the HIV-1_{MAL} genome were required for HIV-1 RNA dimerisation *in vitro* in the absence of NC (Marquet et al., 1991). Dimer formation was found to be temperature-dependent and stimulated by cations.

Purine-rich consensus sequences (RGGARA, where R is a purine) have been identified in the DLS of many retroviruses and provided an explanation for the heterodimer formation and large cation requirements observed (Marquet et al., 1991). However, it was shown that HIV-2 RNA transcripts containing only the 5′ terminal 255 nt of the leader RNA, with no purine-rich motif, could dimerise at 37 °C in high-salt buffer in the absence of Gag protein (Berkhout et al., 1993). Similarly, mutational analysis of HIV-1_{LAI} RNA showed that transcripts that do not contain purine tracts could spontaneously dimerise *in vitro* (Muriaux et al., 1995).

The DLS of MLV and HIV-1, found near the *gag* AUG start codon, were proposed to be multipartite. Supporting this notion, sequences located upstream of the SD and outside of the originally identified DLS were shown to be involved in HIV-1 RNA dimerisation *in vitro* (Marquet et al., 1994). In addition, RNAs lacking most of the HIV-1 genome were shown to form more stable dimers than wild type RNA (Sakuragi and Panganiban, 1997), suggesting the existence of negative regulatory elements.

The primary site of HIV-1_{MAL} RNA dimerisation was subsequently identified using chemical modification interference, which identifies the nucleotides that are required in unmodified form for dimer formation (Skripkin et al., 1994). These nucleotides mapped to a region between the PBS and the SD, and centered around a palindromic sequence located in a hairpin structure. This corresponds to SL-1, as defined in structural studies of the HIV-1 5′ UTR (Figure 1.8A; Clever et al., 1995; Harrison and Lever, 1992). A model was proposed by which dimerisation occured by annealing of the self-complementary sequences, via loop-loop interaction, and the palindrome was therefore named the DIS (Paillart et al., 1994; Skripkin et al., 1994). Thermostability analyses suggested that the sequences 3′ of the SD contributed to the stability of the dimer

(Skripkin et al., 1994), in particular two purine tracts located near the *gag* AUG (Paillart et al., 1994), confirming the multipartite nature of the DLS.

Similar results were obtained with the HIV-1_{LAI} isolate (Clever et al., 1996; Laughrea and Jette, 1994; Muriaux et al., 1995). The core dimerisation domain of HIV-1_{LAI} is located upstream of the SD, in a stem-loop structure containing an exposed palindrome (SL-1) that could initiate dimer formation via loop-loop interaction. The central GC residues of the DIS palindrome were shown to be critical for RNA dimerisation and the presence of two or three flanking adenine residues were found to increase dimer stability (Clever et al., 1996; Paillart et al., 1997). This was confirmed by *in vitro* selection experiments, which also showed that not more than two A-U base-pairs are tolerated in the palindrome and that increasing the length of the palindrome above 6 nt does not increase dimerisation (Lodmell et al., 2000). Similarly, DIS mutants containing four A-U base-pairs were shown to replicate poorly (Laughrea et al., 1999).

Interestingly, duplication of 1000 nt, encompassing the DLS, in an ectopic location of the HIV-1 genome caused the appearance of monomeric genomes in virions, a defect which could be restored by mutation of the authentic DLS (Sakuragi et al., 2001). These results suggested that RNAs containing two DLS self-annealed via intramolecular interactions and that the sequences identified in the 5′ leader RNA are sufficient to promote RNA-RNA interaction.

1.7.2 Formation and maturation of the RNA dimer

- ***The 'kissing-loop' complex model***

The identification of a self-complementary sequence at the centre of the DIS led to the suggestion that dimerisation initiated via symmetrical intermolecular interactions

between theses sequences, forming a so-called 'kissing-loop' complex (Paillart et al., 1994; Skripkin et al., 1994). Supporting this model, mutations preventing self-complementarity in the loop resulted in a dimerisation defect *in vitro* and compensatory mutations restored the process (Clever et al., 1996; Muriaux et al., 1996b; Paillart et al., 1994; Paillart et al., 1996b; Paillart et al., 1997; Skripkin et al., 1994). However, not all mutations that maintain self-complementarity are associated with stable dimer formation *in vitro* (Clever et al., 1996; Laughrea et al., 1999; Paillart et al., 1997), reflecting the fact that only two major DIS variants are observed among natural HIV-1 isolates, out of 64 possible sequences. Further evidence for the kissing-loop model was provided by a mutational analysis of SL-1, demonstrating that the formation of the stem, regardless of its sequence, was required for *in vitro* RNA dimerisation (Clever et al., 1996).

Although the formation of the kissing-loop complex has been well documented and is supported by structural data (Ennifar et al., 2001; Mujeeb et al., 1998), there have been some suggestions that it may not constitute the initiation step of genomic RNA dimerisation. Mutations of the SL-1 palindromic sequence were shown to have no effect on HIV-1_{LAI} genomic RNA dimer formation or stability, but reduced RNA packaging, reverse transcription and viral replication in T-cells (Berkhout and van Wamel, 1996). Interestingly, substitution of SL-1 by a wild type or arbitrary 6 nt palindromic sequence resulted in a replication defect in T-cells but no decrease in dimer formation or stability and replication in PBMCs, suggesting that the requirement for the SL-1 structure is cell type-dependent (Hill et al., 2003). The mobility of the RNA was however different from that of wild type RNA dimers, suggesting that the SL-1 structure is important for the conformation of the dimer.

These discrepancies between different studies might result from the analysis of dimer formation exclusively *in vitro* or in cell culture based-assays. In particular, the latter include cellular and viral factors that may influence the dimerisation process (as discussed below) and contain the full-length genomic RNA, which has been shown to fold into complex structures, possibly regulating RNA dimerisation (Abbink et al., 2005). Interestingly, a recent study of HIV-1 genome dimerisation reported that DIS mutants packaged monomeric genomes that subsequently dimerised, albeit much slower than wild type virus, suggesting that the initiation of dimerisation does occur at the DIS (Song et al., 2007b).

It is worth noting that an alternative mechanism was proposed for the initiation of dimerisation, involving purine-rich sequences and the formation of purine quartets (Awang and Sen, 1993). However, this model was not supported by others and several studies of HIV dimerisation have shown that purine tracts are dispensable for this process (Berkhout et al., 1993; Fu et al., 1994; Muriaux et al., 1995).

- ***Maturation of the RNA dimer***

Transcripts corresponding to the HIV-1_{MAL} and MLV 5′ leader RNA were shown to dimerise *in vitro* in the absence of protein, yet addition of NC (p15) greatly enhanced dimer formation (Darlix et al., 1990; Prats et al., 1990), suggesting that NC played a role in RNA dimerisation. This was confirmed by the observation that a more stable HIV-1 RNA dimer, as determined by its thermal stability, is formed *in vitro* in the presence of NC (Feng et al., 1996b; Muriaux et al., 1996a). This stable dimer is thought to result from the formation of an extended duplex following the initial loop-loop interaction. Despite structural studies supporting the potent formation of an extended duplex by interstrand interactions of residues in the SL-1 stem (Ennifar et al., 1999;

Girard et al., 1999; Mujeeb et al., 1999), it remains uncertain whether an extended duplex exists *in vivo*. Nonetheless, strong evidence supports the transition of a loose dimer into a more stable dimer form, both *in vitro* - in the presence of NC (Feng et al., 1996b; Muriaux et al., 1996a) or at high temperatures (Laughrea and Jette, 1996; Muriaux et al., 1996b) - and in cells (Fu et al., 1994; Fu and Rein, 1993). Importantly, this 'maturation' of the RNA dimer was shown to correlate with proteolytic processing and particle maturation as protease-deficient viruses contained mostly monomeric genomes (Fu et al., 1994; Fu and Rein, 1993; Shehu-Xhilaga et al., 2001a; Song et al., 2007b). In addition, mutations preventing the cleavage at the p2/NC junction inhibit dimer maturation in HIV-1 (Shehu-Xhilaga et al., 2001b), further supporting a role for the proteolytic cleavage in dimer maturation. Interestingly, the Gag-Pol precursor and Pol protein have also been reported to influence the dimer maturation. Increasing the Gag/Gag-Pol ratio by overexpressing PR(-) or PR(+) Gag-Pol expression vectors progressively reduced the thermal stability of the RNA dimer, the overexpression of a PR(-) Gag-Pol vector having a more profound effect on dimer stability (Shehu-Xhilaga et al., 2001a). Furthermore, packaging of Pol is required for dimer maturation, as RNA dimers isolated from Gag particles produced in the absence of a functional Gag-Pol are less stable and show altered mobility compared to wild type dimers (Shehu-Xhilaga et al., 2002). Yet, this effect is not solely caused by the lack of PR. While rescue of PR expression alone is sufficient to restore the stability of the dimer, it does not restore the wild type dimer conformation or correct particle morphogenesis. Co-expression of Vpr-RT or Vpr-IN fusion protein does not rescue dimer conformation; however, supplementation of a Vpr-RT-IN fusion complex in *trans* could restore the conformation of the RNA dimer as well as the particle morphology, demonstrating that the polymerase protein is required for complete dimer maturation and mature core

formation. As a result, one might envisage a physical interaction between the mature RNA dimer and the RT and IN enzymes. As mentioned previously, these components are associated in the PIC, where reverse transcription takes place (see section 1.5.2), and it has been proposed that the dimeric state of the RNA genome plays an important role during reverse transcription and facilitates recombination (Andersen et al., 2003; Balakrishnan et al., 2001; Balakrishnan et al., 2003; Berkhout et al., 1998; Hu and Temin, 1990a; Mikkelsen et al., 2000; Temin, 1991). In HIV-1, the first strand transfer was shown to be more efficient when the RNA template is dimeric (Berkhout et al., 1998). Accordingly, the first strand transfer has been shown to occur in a random manner, with equal contribution from intramolecular and intermolecular transfer (van Wamel and Berkhout, 1998). Furthermore, SL-1 deletion mutants showed a strong reduction in second strand transfer (Paillart et al., 1996a), suggesting that dimeric RNA is required for efficient strand transfer during reverse transcription.

1.7.3 Dimerisation in HIV-2

In HIV-2, two sequences were shown to be involved in RNA dimerisation *in vitro* (Dirac et al., 2001; Jossinet et al., 2001; Lanchy and Lodmell, 2002). Since the purine-rich (PuGGAPuA) motifs initially proposed as dimerisation sites were shown to be dispensable for HIV-2 dimerisation (Berkhout et al., 1993), a self-complementary sequence flanked by unpaired adenines and similar to HIV-1 SL-1 represented a potential dimerisation site. Surprisingly, it was first reported that SL-1 did not mediate dimerisation in HIV-2. Instead, a sequence located in the PBS (see Figure 1.8B) was identified as being the main site of dimerisation and mutational analysis suggested that HIV-2 dimerisation also occurred via formation of a kissing-loop complex (Jossinet et al., 2001). The effect of $tRNA^{lys3}$ annealing on dimerisation was measured and it was

found to inhibit RNA dimerisation, suggesting that the RNA-RNA complex must dissociate to allow tRNA annealing and reverse transcription. However, this is not consistent with the proposed role of dimeric RNA in reverse transcription mentioned above.

Another detailed mutational analysis of the HIV-2 leader showed that SL-1 mediated RNA dimerisation *in vitro* and that the palindromic sequence in the loop (termed DIS and highlighted in red in Figure 1.8B) was essential for dimer formation (Dirac et al., 2001). Interestingly, the authors observed that even though the palindrome could be replaced by other self-complementary sequences, no more than two A-U base-pairs were tolerated, consistent with previous findings in HIV-1 (Laughrea et al., 1999; Lodmell et al., 2000) and suggesting that this might be a common feature of retroviruses. Structural analysis in the context of the entire leader confirmed that the DIS palindrome must be exposed to initiate dimerisation (Figure 1.8B; Dirac et al., 2001, 2002).

In order to determine the exact dimerisation site of HIV-2, Lanchy and Lodmell compared the methods used in the two divergent studies and observed that, *in vitro*, two dimerisation sites can be used under different conditions (Lanchy and Lodmell, 2002). First, a short sequence at the 5′ end of the PBS promotes dimer formation under mild conditions (high potassium chloride, high magnesium dimerisation buffer, electrophoresis at 4 °C in tris-borate-magnesium buffer) and is therefore considered to be a loose-dimer inducing element (Jossinet et al., 2001). Second, a palindrome located within the SL-1 structure, similar to the HIV-1 DIS, promotes the formation of a tight dimer under more stringent conditions which includes a low sodium chloride, high magnesium dimerisation buffer and electrophoresis in tris-borate-EDTA (Dirac et al., 2001). Interestingly, the facility with which one site could substitute for the other

suggests a possible switching mechanism that was postulated to be used in the viral replication cycle (Lanchy and Lodmell, 2002). Further study of the HIV-2 5′ leader RNA suggested that it could adopt two alternative conformations, similar to the LDI and BMH conformations observed in HIV-1, which could regulate RNA dimerisation (Dirac et al., 2002). Dimerisation was proposed to be inhibited when the RNA forms the LDI conformer and facilitated in the BMH conformation. As for HIV-1 (Huthoff and Berkhout, 2001), the NC protein was postulated to promote the LDI-to-BMH switch and to increase dimer formation (Dirac et al., 2002). Another long-distance interaction, between the 5′ UTR and the *gag* AUG start codon, was proposed to negatively regulate RNA dimerisation by silencing SL-1 (Lanchy et al., 2003a; Lanchy et al., 2003b). However, this interaction would be similar to the U5-AUG duplex observed in HIV-1 in the dimerisation-competent BMH conformation (Abbink and Berkhout, 2003; Abbink et al., 2005). Hence, the exact effect of the long-range interaction and conformational switch on HIV-2 dimer formation remains to be confirmed.

Finally, a third palindrome in the HIV-2 Psi region was shown to regulate dimer formation *in vitro* as it could prevent SL-1-mediated dimerisation via SL-1 silencing or promote dimerisation itself (Lanchy et al., 2003a). The location of this dimerisation signal within the packaging determinant (highlighted in yellow in Figure 1.8B) suggested that the two processes may be interconnected.

To date, there has been no confirmation that the DIS or any other palindromic sequence in the HIV-2 5′ leader RNA is sufficient to promote the dimerisation of HIV-2 genomic RNA in the virus.

1.7.4 Relationship between genome dimerisation and encapsidation

As mentioned in section 1.6.1, the HIV-1 *cis*-acting packaging signal overlaps with the core DIS, leading to the suggestion that the two processes are linked and that dimerisation may be necessary for packaging (Paillart et al., 2004b; Russell et al., 2004). Supporting this hypothesis, several studies have reported that mutations which affect dimer formation also reduce genome encapsidation (Berkhout and van Wamel, 1996; Clever and Parslow, 1997; Harrison et al., 1998; Laughrea et al., 1997; McBride and Panganiban, 1996, 1997; Paillart et al., 1996a). Similarly, the HIV-2 packaging signal appeared to overlap with sequences that were important for RNA dimerisation (Dirac et al., 2001; Griffin et al., 2001; Lanchy et al., 2003a; Lanchy and Lodmell, 2002).

Interestingly, structural studies revealed that dimerisation of MLV RNA induces a conformational change, resulting in the exposure of a conserved UCUG motif that bound NC with high affinity (D'Souza and Summers, 2004). Sequestered NC binding sites in the monomeric RNA became exposed upon dimerisation, possibly enhancing the packaging of a dimeric genome. This led the authors to postulate that a structural RNA switch mechanism may be involved in the encapsidation of a diploid genome.

Because of the close proximity of the Psi and DIS elements, it has proven difficult to analyse packaging and dimerisation independently of each other and to determine whether genome dimerisation is indeed a prerequisite for encapsidation. The exact relationship between genome dimerisation and packaging therefore remain poorly understood.

1.8 Aims of research

The aims of this research were to clarify the role of the HIV-2 packaging signal in genome dimerisation and virion infectivity in order to elucidate the relationship between RNA dimerisation, encapsidation, viral replication and particle morphogenesis.

The approach was to use site-directed mutagenesis to mutate the sequence and alter the structure of the HIV-2 packaging and dimerisation signals. The pSVR proviral clone of the HIV-2_{ROD} isolate was used to analyse the mutations in the context of a fully infectious virus. Effects of the mutations on genome encapsidation, RNA dimerisation *in vitro* and *in virio*, protein expression and viral replication were assessed in transiently transfected Cos-1 cells using RNAse protection assay, native northern blot and western blot. Viral infectivity was measured in single- and multiple-round assays using a reporter cell line and T-cells. The effect of the Psi deletion on particle morphology was also evaluated by electron microscopy.

It was hypothesised that identification of the core sequence and structural elements required for HIV-2 genome dimerisation would lead to further investigation into their requirement for RNA packaging and viral infectivity with respect to characterising the relationship between dimerisation and viral replication in HIV-2. In addition, the emergence of potential revertants in the long-term replication study might provide further insight into the elements required for HIV-2 infectivity and indicate new targets to block viral replication.

2 Materials and methods

2.1 Materials

Unless otherwise stated, all chemicals were purchased from Sigma-Aldrich. Ethanol, methanol and isopropanol were purchased from Fisher Scientific. Distilled/deionised water was generated using a Millipore filtration system.

Oligonucleotides were obtained from Sigma-Genosys. Standard oligonucleotides were synthesised at 0.025 μmole and desalted. Mutagenic oligonucleotides used in oligonucleotide-cloning site-directed mutagenesis were synthesised at 0.05 μmole and purified by high performance liquid chromatography (HPLC).

Unless specified in the text, all enzymes, including restriction endonucleases, and their associated buffers were purchased from New England Biolabs (NEB), Roche or Promega.

Commercial reagents and enzymes utilised for RNA work were obtained from Promega or Ambion. Water and prepared reagents required to be RNase-free were treated with 0.1 % (v:v) diethylpyrocarbonate (DEPC) at 37 °C for 12 h prior to autoclaving.

Radiochemicals were purchased from Amersham-GE Healthcare and Perkin Elmer.

Cell culture reagents and media were obtained from Gibco-Invitrogen and PAA. Cell lines already in use in the laboratory were obtained from the American Type Culture Collection (ATCC), the European Collection of Cell Cultures (ECACC) and the Centre for AIDS Reagent (CFAR) at the National Institute for Biological Standards and Controls (NIBSC).

General laboratory plasticware was obtained from Corning, Greiner Bio-One and Starlab. Plastics used in tissue culture were purchased from Corning and Greiner Bio-One.

2.2 Cloning procedures/manipulation of DNA

2.2.1 Digestion of plasmid DNA by restriction endonucleases

Restriction digests were carried out in the appropriate reaction buffer and at the recommended temperature, typically 37 °C, according to the manufacturer's instructions. Bovine serum albumin (BSA) was added to a final concentration of 100 µg/ml when required. Restriction analysis of recombinant plasmid DNA was performed on 1 µg of DNA using 10 Units of enzyme in a total volume of 20 µl and incubated for 1 h. Digestion of DNA subsequently used in ligation was achieved by incubation of 5 µg of DNA and 40 Units of enzyme in a total volume of 50 to 100 µl for 2 h. Simultaneous digestion by *Aat*II and *Xho*I or *Aat*II, *Xho*I and *Xmn*I were carried out in NEB buffer 4 in the presence of 1× BSA.

2.2.2 Dephosphorylation of digested plasmid DNA

Vector plasmid DNA to be used in compatible end ligations was dephosphorylated using 10 Units of Shrimp Alkaline Phosphatase (Roche) for 15 min at 37 °C prior to gel purification.

2.2.3 Polymerase Chain Reaction (PCR)

PCR was performed using either the Taq DNA polymerase from NEB or the BioMix Red reagent from Bioline. Typically, 20 to 100 ng of DNA was used and each primer was added to a final concentration of 0.4 μM.

When the NEB Taq DNA polymerase was used, the standard *Taq* reaction buffer provided was diluted to a final concentration of 1 × (10 mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$, pH 8.3). Each deoxyribonucleotide triphosphate (dNTP; Promega) was added separately to a final concentration of 0.2 mM. When using the 2 × BioMix Red, the DNA and primers were mixed with H_2O and the BioMix Red was added to the final concentration of 1 ×.

Reactions were carried out in a Labnet MultiGene II, a Techne Touchgene or a Perkin Elmer Cetus DNA thermal cycler. When using the latter, the PCR mixture was overlaid with a drop of sterile mineral oil. The hybridisation temperature (T_H) was adjusted for each primer pair as follows: melting temperature (T_M) of the primer (provided by the manufacturer) minus 5 °C. The time of the polymerisation step was calculated as follows: 1 min per kb of DNA to amplify.

A typical PCR amplification comprised a 2 min denaturation step at 95 °C, 30 cycles of 1 min denaturation at 95 °C, 1 min hybridisation at $T_H = T_M - 5$ °C and 1 min/kb polymerisation at 72 °C followed by a final extension of 5 min at 72 °C.

2.2.4 Purification of DNA

Digested vector and insert DNAs used in cloning were separated by agarose gel electrophoresis in 1 × TBE (90 mM Tris-borate, 2 mM EDTA) and purified by gel extraction using the Qiagen "Gel Extraction Kit", according to the manufacturer's instructions.

PCR products were purified using the Qiagen "PCR Purification Kit", according to the manufacturer's instructions.

2.2.5 DNA ligation

Vector and insert DNA were ligated using 2 to 6 Units of T4 DNA ligase (Promega) in 1 × reaction buffer (30 mM Tris-HCl pH 7.8, 10 mM $MgCl_2$, 10 mM DTT and 1 mM ATP). Vector to insert ratio ranged from 1:5 to 1:1 and the final volume of the ligation did not exceed 20 µl. Reactions were incubated at 16 °C overnight and 2 to 5 µl of the reaction was subsequently transformed into a 50 µl aliquot of frozen competent *Escherichia coli* (*E. coli*) bacteria.

2.2.6 Site-directed mutagenesis

Site-directed mutagenesis was performed using the QuickChange II mutagenesis kit (Stratagene) following the manufacturer's instructions. For each reaction, 1 µl of plasmid DNA template (20 to 50 ng) was mixed with 1 µl of each mutagenic primer (10 ng), 1 µl of 10 × reaction buffer, 1 µl of dNTP mix, 4 µl of H_2O and 1 µl of *Pfu* high fidelity polymerase. The PCR-based amplification was programmed as follow: 30 sec at 95 °C, 16 cycles of 30 sec at 95 °C, 1 min at 55 °C and 1 min/kb of template at 68 °C. Samples were subsequently treated with 1 µl of *Dpn*I for 1 h at 37 °C, followed by transformation in XL-10 Gold ultracompetent cells (Stratagene).

2.2.7 Sequencing

All the mutations introduced were verified by sequencing of the plasmid preparation, which was performed off-site by GeneService Ltd. When using plasmid DNA, 100 ng/µl of DNA purified with the Maxiprep or Miniprep commercial kits (Qiagen) was used. When PCR-amplified DNA was used, 10 ng/µl of DNA purified using the Qiagen "PCR Purification Kit" was sent. All primers were diluted to 3.2 µM. Chromatograms were visualised using the Chromas Lite software (www.technelysium.com.au). Analysis and alignments were performed with CLC Free Workbench 4.0 software (www.clcbio.com).

2.3 Plasmid construction and PCR templates

All of the following plasmid constructs generated for the present study had their sequence verified by digestion with restriction endonuclease where possible and by sequence analysis.

2.3.1 Proviral constructs with packaging deletions and substitutions

pSVR is an infectious molecular clone of the ROD strain of HIV-2 (HIV-2_{ROD}; GenBank accession number M15390) (Clavel et al., 1986a; Guyader et al., 1987), which derives from pROD10 (Ryan-Graham and Peden, 1995) and contains a simian virus 40 (SV40) origin of replication inserted in an *Eco*RI site in the vector DNA, 200 bp 5′ of the 5′ LTR (McCann, 1996; McCann and Lever, 1997). The replication characteristics of this clone and proviral construct pSVRDM, containing a 28 nt deletion in the 5' leader between nucleotides 380 and 408 of HIV-2_{ROD}, have been previously described

(Griffin et al., 2001; McCann and Lever, 1997). Restriction sites and nucleotide numbering, where given, are relative to the first nucleotide of the viral RNA genome. Mutations in the 5' leader of HIV-2 were introduced by site-directed mutagenesis into a subclone of HIV-2, pGRAXS (Kaye and Lever, 1998). The mutagenic oligonucleotides used for the introduction of the Pal, SM1, SM2, SM4, SM5, SM6 and SM7 mutations are listed in Table 1. Subclone pGRAXS-SM2 was mutated using oligonucleotides SM1 or SM6 to create pGRAXS-SM3 and pGRAXS-SM8, respectively. Subclone pGRAXS-SM6 was mutated using oligonucleotide SM4 to create pGRAXS-SM9. Sequences from the resulting subclones pGRAXSPal, pGRAXS-SM1, pGRAXS-SM2, pGRAXS-SM3, pGRAXS-SM4, pGRAXS-SM5, pGRAXS-SM6, pGRAXS-SM7, pGRAXS-SM8 and pGRAXS-SM9 were introduced into the provirus by exchanging an *Aat*II (position -1384) - *Xho*I (position 2032) fragment, generating proviral constructs pSVRPal, pSVRSM1, pSVRSM2, pSVRSM3, pSVRSM4, pSVRSM5, pSVRSM6, pSVRSM7, pSVRSM8 and pSVRSM9 respectively.

2.3.2 Envelope-deleted proviral constructs

The envelope-deleted proviral constructs pSVRΔNB and pSVRΔNBDM, which contain a 550 nt deletion in the *env* ORF between position 6369 and 6919, have been described elsewhere (Griffin et al., 2001). pSVRΔNBSM1, pSVRΔNBSM2 and pSVRΔNBSM3 were generated by replacing the *Aat*II (position -1384) - *Xho*I (position 2032) fragment of pSVRΔNB with that of pGRAXS-SM1, pGRAXS-SM2 and pGRAXS-SM3, respectively. A proviral construct carrying a reversion of the DM mutation was also created in order to control for the absence of second-site mutations in pSVRΔNBDM which might have been introduced during the cloning process. The pSVRΔNBDM

plasmid was mutated by site-directed mutagenesis using mutagenic oligonucleotide DM Rev (Table 1), creating construct pSVRΔNBDMrev.

Table 1: Mutagenic oligonucleotides used to introduce packaging signal deletion and substitutions by site-directed mutagenesis.

Mutagenic oligonucleotides	
Name	Sequence
Pal	5' GAACAAACCACGACTAGAAAGGCGCGGG 3'
SM1	5' GGCGCGGGCCGACCAACCAAAGGCAGCGTGTGG 3'
SM2	5' GGAACAAACCACGACCCTCTGCTCCTAGAAAGGCG 3'
SM4	5' CAAACCACGACGGAGTGGAGGTAGAAAGGCGCGGG 3'
SM5	5' CAAACCACGACCCTCGTGAGGTAGAAAGGCGCGGG 3'
SM6	5' CCAAAGGCAGCGTGTCCTCCGGGAGGAGAAGAGG 3'
SM7	5' CGACGGAGTGCTCCACACGCGGGCCGAGGTAC 3'
DM Rev	5' GGGCGGCAGGAACAAACCACGACGGAGTGCTCCTAGAAAGGCGCGG GCCGAG 3'

2.3.3 Envelope expressor constructs

pCMV-VSVG was a gift from Nijsje Dorman (Dorman, 2000) and contains the coding sequence of the vesicular stomatitis virus (VSV) G glycoprotein in the context of the pCDNA3 expression vector (Invitrogen). Expression of the VSV-G protein is controlled by the hCMV immediate early promoter.

2.3.4 Constructs for protein expression

The pGEXGag2 plasmid used for the expression of the Glutathione-S-transferase (GST) tagged HIV-2 Gag protein in *E. coli* bacteria culture has been described previously (Myers and Allen, 2002).

2.3.5 Constructs used as templates for riboprobes in northern blot and RNase protection assay

Plasmids KS2ΨKE and KS2ΨEP have been previously described and contain HIV-2 sequences from positions 306 to 751 and (-107) to 306 of pSVR, respectively, cloned into the polylinker of the pBluescript KSII(+) transcription vector (Stratagene) (Griffin et al., 2001; Kaye and Lever, 1998).

A glyceraldehyde-3-phosphate dehydrogenase (GAPDH) riboprobe template was created by reverse transcription-PCR (RT-PCR) amplification of the sequence from position 61 to 346 of the human GAPDH using 5' GGTGAAGGTCGGAGTCAACG 3' and 5' AATTAACCCTCACTAAAGGACTCCACGACGTACTC 3' primers. Total cytoplasmic RNA extracted from Jurkat T-cells served as template for the reaction.

2.3.6 PCR-derived templates for *in vitro* transcription

For the *in vitro* dimerisation assay, WT and mutant DNA templates were produced by PCR amplification of the corresponding proviral plasmid using HIV-2 Psi F 5' TAATACGACTCACTATAGGCTGAGTGAAGGC 3' and Psi R 5' AGGTACTT-ACCTTCACCC 3'.

2.4 Preparation of plasmid DNA from bacterial cultures

2.4.1 Strains of Escherichia coli

- ***DH5α***

For general cloning, the DH5α strain of *E. coli* [F^- φ80dlacZΔM15 Δ(lacZYA-argF)U169 deoR recA1 endA1 hsdR17(rk^-, mk^+) phoA supE44 λ^- thi-1 gyrA96 relA1] was used. Competent DH5α were prepared as described in section 2.4.2 from an original stock of frozen chemically competent bacteria purchased from Invitrogen.

- ***Top10F'***

For the transformation of proviral plasmid DNA, the Top10F′ strain of *E. coli* [F′ (lacIq Tn10(TetR)) mcrA Δ(mrr-hsdRMS-mcrBC) φ80dlacZΔM15 ΔlacX74 deoR recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG] (Invitrogen) was used. Top10F′ *E. coli* are deficient for Recombinase A1 (recA1) to reduce occurrence of non-specific recombination in cloned DNA and contain an F′ episome, which carries the tetracycline resistance gene and allows isolation of single-stranded DNA from vectors that have an f1 origin of replication. Chemically competent Top10F′ *E. coli* were obtained from Invitrogen.

2.4.2 Preparation of frozen competent *E. coli*

Frozen competent bacteria were prepared based on the method developed by Inoue (Inoue et al., 1990). Using a sterile loop, competent DH5α (Invitrogen) were plated on Luria Bertani (LB)-agar (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl, pH 7.0, 15 g/l agar) and grown at 37 °C overnight. A single colony was picked with a sterile toothpick,

resuspended in 10 ml of SOB medium (20 g/l tryptone, 5 g/l yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, pH 7.0) and grown overnight at 37 °C, 225 rpm in a shaking incubator. 500 µl of the overnight culture was used to inoculate 250 ml of SOB medium and the bacteria were grown at room temperature (rt) with gentle rotation until an optical density (OD) at 600 nm of 0.6 was obtained. The bacteria were then incubated on ice for 10 min followed by pelleting at 2500 × g at 4 °C for 10 min in a Sorvall RC3C-Plus centrifuge. The pellet was gently resuspended in 80 ml ice-cold TB (10 mM PIPES, 15 mM $CaCl_2$, 250 mM KCl, pH adjusted to 6.7, then $MnCl_2$ added to 55 mM). The suspension was incubated on ice for 10 min followed by centrifugation as above. The pellet was resuspended in 20 ml ice-cold TB supplemented with 7 % (v:v) dimethyl sulfoxide (DMSO), swirled gently and incubated on ice for 10 min. The bacteria were frozen in liquid nitrogen and stored in 50 µl aliquots at -80 °C.

2.4.3 Transformation of competent *E. coli*

A frozen aliquot of competent *E. coli* was thawed on ice for 5 to 10 min. Approximately 100 ng of plasmid DNA was added to the cells. The reaction was mixed by taping the tube gently and immediately placed back on ice. The cells were incubated on ice for 30 min to 1 h before being heat shocked for 45 sec at 42 °C. The cells were put back on ice for 2 min before addition of 500 µl preheated LB medium (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl, pH 7.0). The cells were allowed to recover at 37 °C for 1 h with shaking at 200 rpm. The cells were then spun at 1000 × g for 5 min in a benchtop microcentrifuge and resuspended in 100 µl LB medium prior to plating on LB-agar containing 100 µl/ml ampicillin (Sigma-Aldrich). Plates were incubated at 37 °C for 24 h. When Top10F′ bacteria were transformed with proviral plasmids, plates were incubated at 30 °C for 24 to 36 h to reduce the risk of recombination.

2.4.4 Growth of bacterial culture

All strains of *E. coli* transformed with standard plasmids carrying the Ampicillin resistance (*amp*^{R}) gene were grown in LB supplemented with 100 µg/ml ampicillin. 5 ml cultures were grown in 30 ml universal tubes for small scale preparations of plasmid DNA and 500 ml cultures were grown in 2 l Erlenmeyer flask for large scale preparations. DH5α *E. coli* were incubated at 37 °C whereas Top10F′ *E. coli* transformed with proviral constructs were incubated at 30 °C to prevent recombination.

2.4.5 Small scale preparation of plasmid DNA

A single colony was picked from a plate of transformed *E. coli* (see section 2.4.3) and resuspended in 5 ml of LB containing 100 µg/ml ampicillin (LB-Amp). The culture was incubated at 37 °C, 225 rpm overnight. If Top10F′ transformed with proviral DNA were used, the incubation temperature was decreased to 30 °C. 1 ml of the culture was then spun at 9500 × g for 5-10 min in a Heraeus Biofuge *pico* microcentrifuge and plasmid DNA was extracted by the alkaline lysis method using the Qiagen "Qiaspin Miniprep" kit. The manufacturer's instructions were followed with the exception that plasmid DNA was resuspended in water instead of elution buffer EB which contains EDTA and may inhibit subsequent enzymatic reactions.

Integrity of the plasmid DNA was verified by agarose gel electrophoresis before either sequencing or large scale preparation. Where possible, a test restriction digest was performed to analyse the presence of the introduced mutation.

2.4.6 Large scale preparation of plasmid DNA

250 to 500 µl of the 5 ml culture used for small scale preparation was added to 250 to 500 ml of LB-Amp in a 2 L Erlenmeyer flask and incubated at 37 °C (or 30 °C for Top10F'), 225 rpm overnight. The culture was spun at 4500 × g for 20 min at 4 °C in a Sorval Evolution *RC* centrifuge (rotor SLA-3000). Plasmid DNA was prepared by the alkaline lysis method using the "Qiagen Maxiprep" kit following the manufacturer's instruction with the exception that plasmid DNA was resuspended in 500 µl water instead of buffer EB.

Integrity of the plasmid DNA was verified by agarose gel electrophoresis. The concentration and the level of proteins in the sample were determined by measuring the OD at 260 nm and the OD_{260nm}/OD_{280nm} ratio using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher), respectively.

All plasmids prepared by Maxiprep had their sequence verified (see section 2.2.7).

2.5 Cell culture techniques

2.5.1 Cell lines

- ***Cos-1 cells***

Cos-1 cells were obtained from the ECACC (catalogue number: 88031701). Cos-1 cells are derived from African Green Monkey kidney fibroblasts (CV-1 cells) transformed with an origin-defective mutant of Simian virus 40 (SV40) which encodes the wild type T antigen (Gluzman, 1981). Cos-1 cells were maintained in complete Dulbecco's modified Eagle's medium (DMEM; see section 2.5.2). DEAE-dextran transfection of pSVR-derived proviral DNA (section 2.5.7), which carries a SV40 origin of replication,

into Cos-1 cells yielded a high level of HIV-2 particle release as measured by a reverse transcriptase (RT) assay (see section 2.8.5).

- ***GHOST indicator cell line***

The GHOST indicator cells, engineered by Drs. Littman and Kewal Ramani (Trkola et al., 1998), were obtained from the CFAR at NIBSC (catalogue number: ARP084). GHOST cells are human osteosarcoma (HOS) CD4 positive cells expressing the CCR5 and CXCR4 chemokines on their surface and containing an HIV-2 LTR linked to a green fluorescent protein (*gfp*) gene, each under a specific selection (see section 2.5.3). GHOST cells were maintained in complete DMEM supplemented with antibiotics (see sections 2.5.2 and 2.5.3). Upon infection with HIV-2, the Tat protein *trans*-activates the HIV-2 promoter and triggers the expression of GFP, which can be detected by fluorescence activated cell sorting (FACS) as described in section 2.6.2.

- ***C33-A cells***

C33-A cells, a human cervical carcinoma cell line (Auersperg, 1964), were obtained from the ATCC (catalogue number HTB-31) and were tested negative for human papilloma virus (Yee et al., 1985). C33-A cells were maintained in complete DMEM (see section 2.5.2) and transfected using the Calcium-phosphate method (see section 2.5.7).

- ***Jurkat cells***

Jurkat cells (Jurkat Clone E6-1) were obtained from the CFAR at NIBSC (catalogue number: ARP027). Jurkat cells are a CD4 positive human T-cell leukemia cell line expressing high levels of interleukin-2 (IL-2) after stimulation (Weiss et al., 1984). Jurkat cells were maintained in complete Roswell Park Memorial Institute media (RPMI-1640; see section 2.5.2). HIV-1 and -2 can establish persistent infection in Jurkat cells, with sustained levels of virus replication for prolonged periods.

- ***PM1 cells***

PM1 cells were a gift from Dr. Truus Abbink and can be obtained from the CFAR at NIBSC (catalogue number: ARP057). PM1 cells are a clonal derivative of HUT78, a human cutaneous T-cell lymphoma from peripheral blood of a patient with Sezary Syndrome (Lusso et al., 1995). PM1 cells were maintained in complete RPMI-1640 (see section 2.5.2). PM1 cells have the unique characteristic of supporting infection by a wide range of HIV-1 isolates, including primary and biologically pure macrophage-tropic isolates (Lusso et al., 1995).

- ***SUP-T1 cells***

SUP-T1 cells were a gift from Dr. Truus Abbink and can be obtained from the CFAR at NIBSC (catalogue number: ARP024). SUP-T1 cells are Non-Hodgkin's T-cell lymphoma which express high levels of CD4 (Smith et al., 1984). SUP-T1 cells were maintained in complete RPMI-1640 (see section 2.5.2).

2.5.2 Passage of cell lines

All media were completed with 10 % (v:v) fetal bovine serum (FBS; Gibco BRL) and 1 % penicillin-streptomycin solution (10000 U/ml penicillin, 10 mg/ml streptomycin; Gibco BRL). All cells were grown in humid incubators at 37 °C in the presence of 5 % CO_2.

Cos-1 and GHOST cells were grown in complete DMEM (Gibco BRL) to 80 % confluency. Additional antibiotics, as detailed in section 2.5.3, were used to passage the GHOST cells. Cells were split 1/10 twice a week. Cells grown in 75 cm^2 flasks were washed in pre-warmed phosphate buffered saline (PbS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4) and detached by addition of 2.5 ml 1 × trypsin-EDTA solution (Gibco BRL). Trypsin was inactivated by addition of 7.5 ml serum-containing DMEM before 1 ml of cells was transferred to a new 75 cm^2 flask.

Jurkat, PM1 and SUP-T1 cells were grown in complete RPMI-1640 (Gibco BRL) and split 1/10 with fresh media twice a week.

2.5.3 Antibiotic selection

GHOST cells were maintained and stored under selection for the expression of CD4, co-receptors (CCR5 and CXCR4) and the HIV-2 LTR controlling expression of the *gfp* gene by supplementing the DMEM with 100 µg/ml hygromycin (Invitrogen), 500 µg/ml geneticin (PAA) and 1 µg/ml puromycin (Sigma-Aldrich).

2.5.4 Cell viability and counting

A 10 µl aliquot was removed from thoroughly mixed cell suspensions and added to 10 µl of a trypan blue solution (Sigma-Aldrich). The solution was homogenized by

repeated pipetting prior to loading of 10 µl onto each chamber of an improved Neubauer haemocytometer. An average of two counts of the appropriate area of the grid was used to determine the cell density.

2.5.5 Freezing of cell stocks

Adherent cells were washed in PbS and detached by addition of 1 × trypsin-EDTA for 5 min. Trypsin was inactivated by fresh serum-containing media and the cells were pelleted by centrifugation at 300 × g for 5 min in a MSE Falcon 6/300 centrifuge. Suspension cells were pelleted as before directly from the medium. Cell pellets were resuspended in 10 ml of the appropriate media supplemented with 20 % FBS and 10 % DMSO and 1 ml aliquots were placed at -80 °C overnight prior to storage in liquid nitrogen to allow for slow cooling down of the cells.

2.5.6 Resuscitation of cell stocks

Cells that were stored in liquid nitrogen were thawed at rt and added to 10 ml of pre-warmed medium. Residual DMSO was removed by centrifugation at 300 × g for 5 min. The pelleted cells were resuspended in 10 ml of fresh pre-warmed media and incubated in a 25 cm^2 flask at 37 °C, 5 % CO_2.

2.5.7 Transfection techniques

- ***DEAE-Dextran method***

For the transfection of Cos-1 cells, the DEAE-dextran method was adapted from Mortlock (Mortlock et al., 1993). The day prior to transfection, 1.5 × 10^6 Cos-1 cells

were seeded into a 10 cm dish in fresh DMEM. The following day, 5 to 10 µg of plasmid DNA were mixed with 1.9 ml of pre-warmed PbS and 100 µl of DEAE-dextran (10 mg/ml in 1M Tris-HCl pH 7.5). The cells were washed twice with PbS before the DNA mixture was added. Cells were incubated for 30 min at 37 °C, 5 % CO_2. The cells were then covered with 5 ml of 80 µM chloroquine diphosphate (prepared in serum-free DMEM media) prior to a further 2.5 h at 37 °C. The mixture was removed and the cells were shocked for 2 min with serum-free DMEM containing 10 % DMSO. Following two washes in serum-free DMEM, the cells were incubated in 10 ml of fresh complete DMEM for 48 h at 37 °C, 5 % CO_2.

- ***Calcium-phosphate method***

C33-A cells were transfected using the calcium-phosphate method. The day prior to transfection, 1.5×10^6 C33-A cells were seeded into a 10 cm dish in fresh complete DMEM. The following day, 30 µg of plasmid DNA were diluted in 645 µl of H_2O and mixed with 733 µl of 2 × HBS (50 mM HEPES pH 7.1, 250 mM NaCl, 1.5 mM Na_2HPO_4) in a glass vial. 88 µl of 2M $CaCl_2$ were added and the mixture was incubated at room temperature for 20 min. The mixture was then added dropwise to the cells and these were incubated at 37 °C for 24 h. After 24 h, the culture medium was removed, the cells were washed twice with warm PbS and fresh complete DMEM was added before incubation at 37 °C for a further 24 h.

2.6 Virus culture and analysis

2.6.1 Preparation of concentrated virus supernatants

Virus-containing supernatant from transfected or infected cells was collected and filtered through a 0.45 µm filter. If the virus was isolated from suspension cells, the supernatant was first spun at 300 × g for 5 min to pellet the cells. Filtered viruses were precipitated in 0.5 V of 30 % (w:v) polyethylene glycol (PEG)-8000 (MW) in 0.4 M NaCl overnight at 4 °C. Precipitates were spun at 500 × g for 40 min at 4 °C and the supernatant was discarded. Following a second 5 min spin to remove any traces of PEG, the virus-containing pellet was resuspended in 500 µl of TNE (10 mM Tris-HCl pH 7, 150 mM NaCl, 1 mM EDTA pH 8) and layered onto 500 µl of TNE-20 % (w:v) sucrose. The virus preparation was spun at 71550 × g in a Beckman Optima MAX Ultracentrifuge (Rotor TLA-55) for 2 h at 4 °C. The supernatant was removed and the virus was resuspended in the appropriate medium depending on the downstream application (see sections 2.6.2, 2.6.3, 2.6.5 and 2.7.5).

The few remaining microliters of virus-containing TNE were used to determine the RT activity of the preparation as described in section 2.8.5.

2.6.2 Infectivity assay using the GHOST cells

For each virus to be analysed, two 10 cm dishes of Cos-1 or C33-A cells (1.5×10^6 cells per dish) were transfected with 5 µg of proviral DNA as described in section 2.5.7 and incubated for 48 h.

The day prior to infection, 2×10^5 GHOST cells were seeded in 6-well plates in selection medium (see section 2.5.3).

At 48 h post-transfection, the viruses were concentrated as described above, resuspended in 500 µl DMEM supplemented with 1 × penicillin/streptomycin, 10 % FBS and 8 µg/ml polybrene (hexadimethrine bromide, Sigma-Aldrich) and pooled. The RT activity of each virus preparation was determined as described in section 2.8.5 and used to calculate the virus input. Increasing amounts of virus were used to infect the GHOST cells in a total volume of 1 ml DMEM supplemented with 1 × penicillin/streptomycin, 10 % FBS and 8 µg/ml polybrene. The cells were incubated at 37 °C, 5 % CO_2 for 16 h. The medium was then replaced with fresh DMEM supplemented with 1 × penicillin/streptomycin, 10 % FBS, hygromycin (100 µg/ml), geneticin (500 µg/ml) and puromycin (1 µg/ml) and the cells were incubated at 37 °C, 5 % CO_2 for a further 48 h.

The cells were next washed with PbS and detached using 400 µl of 1 × trypsin-EDTA prior to addition of 600 µl of DMEM. The cells were spun at 300 × g for 5 min and resuspended in 200 µl of a freshly prepared 2 % (w:v) paraformaldehyde (Sigma-Aldrich) solution to fix the cells and inactivate the virus. The expression of GFP was measured by FACS using a BD FACSCalibur system (BD Biosciences) and the results were analysed with the WinMDI software (http://facs.scripps.edu/software.html).

2.6.3 T-cell replication assay

For the long-term study of viral replication, Jurkat T-cells were infected with HIV-2 produced from transfection of Cos-1 cells and passaged over a period of approximately 3 months. For each virus to be analysed, two 10 cm dishes of Cos-1 (1.5×10^6 cells per dish) were transfected with 5 µg of proviral DNA as described in section 2.5.7 and incubated for 48 h. The viruses were concentrated as described in section 2.6.1, resuspended in 500 µl complete RPMI-1640 and pooled. The RT activity of each virus

preparation was determined as described in section 2.8.5 and used to calculate the virus input. Typically, an equivalent of 10000 cpm was used to infect 1×10^6 Jurkat cells in 2 ml complete RPMI-1640 supplemented with 10 µg/ml DEAE-dextran. The cells were incubated overnight at 37 °C, 5 % CO_2 before being transferred into a 25 cm^2 flask containing 8 ml of complete RPMI-1640. The cells were returned to the incubator for a further 4 to 6 days, at which point the first sample was taken. Two 10 µl samples were used to determine the RT activity (see section 2.8.5). The cells were then split ½ every 8 days and viral replication was followed by measuring the RT activity of the culture every 4 days.

For the short-term study of viral replication and the analysis of revertants, SUP-T1 and PM1 cells, which produce a higher titre of HIV-2 than Jurkat cells, were used. Viruses were produced by transfection of Cos-1 cells (section 2.5.7) and purified as described in section 2.6.1. The RT activity of each virus preparation was determined as described in section 2.8.5 and used to calculate the virus input required. Typically, 5000 to 10000 cpm was used to infect 1×10^6 PM1 or SUP-T1 cells in a total of 5 ml complete RPMI-1640. Cells were maintained at 37 °C, 5 % CO_2 and split 1/5 twice a week. Viral replication was followed by measuring the RT activity of two 10 µl samples every 2 to 4 days.

2.6.4 Passage of viruses

Viruses cultured in PM1 cells were passaged in order to study the evolution of the mutant viruses. Following infection of PM1 cells as described above and once the peak of infection was reached as determined by RT activity, 500 µl of virus-containing cell-free supernatant was used to infect 1×10^6 PM1 cells in 5 ml RPMI-1640 supplemented with antibiotics and FBS. The same aliquot of virus was frozen at -20 °C and the cell

pellet was used for sequencing analysis. The cell pellet was resuspended in 500 µl of PbS followed by addition of 100 µl of lysis buffer (10 mM Tris-HCl pH 8, 1 mM EDTA, 0.5 % Tween-20) and 6 µl of 20 mg/ml Proteinase K (Ambion). The mixture was mixed gently and incubated at 56 °C for 1 h to inactivate any residual virus. After inactivation of the proteinase K for 10 min at 95 °C, 1 µl of the genomic DNA preparation was used in a 50 µl PCR using different sets of HIV-2 specific primers (see Table 2) to amplify the leader region (1F and 751R), the *gag* ORF (514F and 1440R; 1221F and 2144R; 1800F and 2627R) and the U3 promoter region (NEF2F and 300R; RODU3 and 300R; -107F and 300R). Amplifications were performed as described previously (section 2.2.3) except that the number of cycles was increased to 40 due to the low amount of template present in the sample. The PCR amplicons were purified using the Qiagen "PCR purification kit" and sent for sequencing in order to determine if reversion had occurred.

2.6.5 Electron microscopy of purified virus particles

Envelope-deleted viruses were produced by transfection of pSVRΔNB and pSVRΔNBDM into Cos-1 cells as described in section 2.5.7. Viruses were purified on a 20 % sucrose-cushion as detailed above and washed three times in 50 µl 100 mM NaCl, 50 mM MOPS (pH 7) buffer by centrifugation at 6000 × g for 8 min at 4 °C. Pelleted virions were resuspended in 10 µl of the same buffer at 4 °C for 24 h prior to negative staining. 0.8 µl of virion preparation was applied to a carbon coated grid and stained with a few drops of 1 % uranyl acetate. Micrographs were recorded on a Philips EM208S at a nominal magnification of × 20,000. Electron microscopy was performed in the Structural Studies division at the Laboratory of Molecular Biology, in collaboration with R.A. Crowther.

Table 2: Sequences of HIV-2 specific oligonucleotides used for the sequencing of potential revertants in the evolution study.

Primer names	Primer sequences
1F	5' GGTCGCTCTGCGGAGAGGCTGG 3'
751R	5' GGTACCATTGGATCTAAAACTG 3'
514F	5' CCTACCTTTAGACAGGTAG 3'
1440R	5' CATAGCTTTGGAACGGC 3'
1221F	5' GCGGGGCAGCTTAGAGAGCC 3'
2144R	5' CTACTGGCTGACCCTC 3'
1800F	5' GGACACATCATGACAAACTGC 3'
2627R	CCTTGTTTAGTTCTCTG 3'
NEF2F	5' GTGGAGGTATTGGAACG 3'
RODU3	5' GGAAGGGATGTTTTACAG 3'
-107F	5' CTGCAGGGACTTTCCAGAAGG 3'
300R	5' TCCCTGTTCAGGCGCCAACC 3'

2.7 Synthesis, preparation and analysis of RNA

2.7.1 *In vitro* transcription

RNA used in the *in vitro* dimerisation assays was produced as follows. 5 µl of PCR-derived template (see section 2.3.6) was mixed with 5 µl of 10 × transcription buffer (Ambion), 1.5 µl RNasin (40 U/µl; Promega), 2 µl each of 10 mM ATP, GTP and CTP (Promega), 6 µl of 100 µM UTP (Promega), 2 µl [α-^{32}P] UTP 800 Ci/mmol (20 mCi/ml; GE Healthcare or Perkin Elmer), 22 µl H_2O and 2.5 µl T7 RNA Polymerase (Ambion). The reaction was incubated at 37 °C for 1 to 2 h before addition of 2 µl of TURBO DNase (Ambion) and incubation at 37 °C for 15 to 30 min. RNAs were purified as described in section 2.7.2.

Antisense riboprobes used in the ribonuclease protection assays (RPAs) were synthesised using the T3 RNA polymerase (Promega). 2 µl of linearised template (0.5 mg/ml) was mixed with 4 µl of 5 × transcription buffer (Promega), 2 µl of 100 mM DTT, 0.5 µl RNasin (40 U/µl), 1 µl of each 10 mM ATP, GTP and CTP, 2.4 µl of 100 µM UTP, 2.5 µl [α-^{32}P] UTP 800 Ci/mmol (20 mCi/ml), 1 µl of T3 RNA Polymerase and 2.6 µl of H_2O. The reaction was incubated at 37 °C for 1 to 2 h before addition of 1 µl of TURBO DNase and incubation at 37 °C for 15 min. 21 µl of 2 × loading buffer II (Ambion) were added and the probes were gel-purified as described below.

Antisense riboprobes used in the northern blot were synthesised as follow. 10 µl of linearised DNA template (0.5 mg/ml) was mixed with 5 µl of 10 × Biotin RNA labelling mix (Roche), 10 µl of 5 × transcription buffer (Promega), 5 µl of 100mM DTT, 1.5 µl of RNasin (40 U/µl), 13.5 µl of H_2O and 5 µl of T3 RNA Polymerase (Promega) and incubated at 37 °C for 1 h. 5 µl of RQ DNase I (Promega) were added and the reaction was incubated at 37 °C for a further 15 min. Following purification through a G-50 sephadex column, riboprobes were aliquoted and stored at -20 °C.

2.7.2 Purification of *in vitro* transcribed RNA

Radiolabelled RNAs used in the *in vitro* dimerisation assay were quantified by TCA precipitation (see section 2.7.3) and purified as follow. One volume of 25:24:1 (v:v:v) phenol:chloroform:isoamyl alcohol (IAA) at pH 4.7 (Sigma-Aldrich) was added to the transcription reaction, thoroughly mixed by vortexing and spun for 2 min at 16000 × g. The top layer was transferred into a clean tube containing one volume of 24:1 (v:v) chloroform-IAA (Sigma-Aldrich), mixed by vortexing and spun for 2 min at 16000 × g. The top layer was transferred into a clean tube containing 0.1 V 5 M NH_4OAc and 2.5 V 100 % EtOH, mixed by vortexing and precipitated at - 80 °C for at least 30 min. The

RNA was pelleted by centrifugation at 16000 × g for 15 min at 4 °C, and resuspended in the appropriate volume of DEPC-treated water.

Riboprobes used in the RPA were gel-purified on an 8 M urea 5 % acrylamide gel run in 1 × TBE at 250 Volts for 30 min. After electrophoresis, the gel was wrapped and exposed to X-ray film for 1 to 3 min. The riboprobe bands were cut out and eluted in 350 µl of probe elution buffer (Ambion) at 37 °C overnight.

Riboprobes used in the northern blot were purified through a G-50 sephadex column (Roche) following the manufacturer's instructions.

2.7.3 TCA precipitation

Radiolabelled RNAs used in *in vitro* dimerisation assays were quantified by the TCA precipitation method. After incubation with DNase, 1 µl of radiolabelled RNA was removed from the reaction and diluted in 10 µl of H_2O. 1 µl of the dilution was mixed with 10 µl of 10 mg/ml yeast tRNA, 89 µl of H_2O and 500 µl of 5 % trichloroacetic acid (TCA) and incubated on ice for 10 min. The reaction mixture was passed onto a glass microfibre filter previously equilibrated with 5 % TCA. The filter was washed once with 5 % TCA, once with acetone and air dried. As a control for incorporation, 1 µl of the original dilution was spotted onto another filter and air dried. Both filters were placed into scintillation vials, covered with 1 ml of OptiPhase HiSafe 2 scintillant and the percentage of incorporation was determined using a TRI-CARB 2100 TR liquid scintillation analyser (Packard).

2.7.4 Isolation of cytoplasmic RNA

Cytoplasmic RNA was extracted using the Qiagen "RNeasy mini kit".

Adherent cells were washed once in cold PbS, lysed directly in the dish by addition of 600 µl of buffer RLT from the "RNeasy mini kit" and the cytoplasmic RNA was extracted following the manufacturer's instructions.

If suspension cells were used, the cells were pelleted by centrifugation at 300 × g for 5 min, washed in cold PbS once and the cell pellet was resuspended in 600 µl of buffer RLT. The cytoplasmic RNA was subsequently extracted following the manufacturer's instructions and eluted in 40 to 50 µl of RNase-free H_2O.

2.7.5 Isolation of virion RNA

Virion RNA was extracted using the Qiagen "QiAamp Viral RNA mini Kit", following the manufacturer's instructions. Virions were pelleted as described in section 2.6.1 and resuspended in 140 µl of RPMI-1640 prior to extraction. The RNA was eluted in 40 to 50 µl of buffer AVE.

2.7.6 DNase treatment of RNA samples used in the RPA

RNA samples used in the RPA were DNase treated to remove any trace of plasmid DNA carried over during the purification process. Cytoplasmic and virion RNA samples purified using commercial kits were eluted in 40 µl of H_2O and buffer AVE, respectively. 5 µl of 10 × TURBO buffer and 5 µl TURBO DNase (Ambion) were added and the reactions were incubated at 37 °C for 30 min. RNAs were then extracted once with phenol-chloroform and once with chloroform as described above and precipitated with 0.1 V 5 M NH_4OAc and 2.5 V 100 % EtOH at -80 °C for at least 30 min. Samples were spun at 16000 × g for 15 min and resuspended in 50 µl of RNase-free H_2O.

2.7.7 RNase Protection Assay

RPAs were performed using the "RPA III kit" from Ambion as recommended by the manufacturer. Riboprobes were synthesised and purified as described in sections 2.7.1 and 2.7.2. Following overnight elution of the probe in the probe elution buffer (Ambion), a 10 µl aliquot was removed, added to 2 ml OptiPhase HiSafe 2 scintillant and counted in the liquid scintillation analyser (Packard). 2×10^5 cpm was used for each riboprobe tested. Cytoplasmic RNA inputs were normalised on concentration, typically 0.5 µg. Virion RNA inputs were normalised on RT activity as determined by a RT assay (see section 2.8.5), with an equivalent of 2500 cpm being used. Cytoplasmic and virion RNAs were first co-precipitated with the KS2ΨKE, KS2ΨEP or GAPDH riboprobes. The probes were allowed to anneal at 42 °C overnight prior to digestion of single stranded RNA using a 1:100 dilution of RNaseA/T1 provided in the kit for 30 min at 37 °C. The protected RNA fragments were precipitated using the RNase inactivation/precipitation solution, resuspended in 10 µl of loading buffer II (Ambion) and separated on an 8 M urea, 5 % polyacrylamide gel alongside radiolabelled "Century Plus" RNA size markers (Ambion). The gels were run at 1100 Volts for approximately 2 h, dried and autoradiographed. RNAs were quantified using an Instant Imager (Packard) or by densitometry using the ImageJ software (http://rsb.info.nih.gov/ij/). The GAPDH probe was used to correct the level of cytoplasmic RNA for loading variations, with the mock transfected lane set at 100 %. Plasmid DNA contamination was accounted for using the KS2ΨEP probe, which allows differential detection of DNA (protected fragment of 413 nt) and RNA (protected fragment of 306 nt). The packaging efficiency was determined as follows: $[(\text{viral RNA})/(\text{cyto RNA})]_{\text{Mutant}}$ / $[(\text{viral RNA})/(\text{cyto RNA})]_{\text{WT}}$.

2.7.8 Northern blot

Northern blots were performed using the "NorthernMax-Gly kit" and "BrightStar Biodetect kit" from Ambion, following the manufacturer's instructions.

Cytoplasmic and virion RNA samples were normalised on concentration and RT activity as before, with a typical amount of 2 µg cytoplasmic RNA and an equivalent of 2500 cpm virion RNA used. The samples were electrophoresed at room temperature on a non-denaturing 0.8 % agarose-LE (Ambion) gel in 1 × gel prep/running buffer (Ambion) for 4 h at 60 to 80 Volts. The Biotinylated Millenium RNA size markers (Ambion) were run alongside the samples. The RNA was transferred onto a positively charged nylon membrane (BrightStar membrane; Ambion) overnight using the supplied transfer buffer. The membranes were cross-linked by baking at 80 °C for 15 min and pre-hybridised in UltraHyb for 30 min at 68 °C. Biotinylated KS2ΨKE (0.4 nM final concentration) was hybridised overnight at 68 °C in UltraHyb. The membrane was washed once in 2 × standard sodium citrate (300 mM NaCl, 30 mM sodium citrate, pH 7; SSC), 0.1 % SDS for 10 min at rt and twice in 0.1 × SSC, 0.1 % SDS for 15 min at 68 °C. The blot was detected using the BrightStar Biodetect kit from Ambion following the manufacturer's protocol. The blot was analysed by densitometry using the ImageJ software and the percentage of dimer was calculated.

2.7.9 *In vitro* dimerisation assay

The *in vitro* dimerisation assay was adapted from the method previously described by Dirac and co-authors (Dirac et al., 2001). 20 to 40 nM of radiolabelled Psi RNAs (position 350 to 450) were incubated in 10 µl of dimer buffer (10 mM Tris-HCl pH 7.5, 40 mM NaCl and 5 mM $MgCl_2$) for 10 min at 65 °C, followed by slow cool down to room temperature. Dimers were stabilised on ice for 2 min and 2 µl of native loading

buffer containing 30 % glycerol were added. Samples were analysed on a 1.5 % agarose gel run in 0.25 × TBE for 1 h 30 at room temperature at 30 to 40 mAmp. Gels were fixed in 40 % MeOH, 10 % acetic acid, dried and autoradiographed. RNA was quantified by densitometry using the ImageJ software.

2.7.10 RT-PCR

For the amplification of viral sequences from the HIV-2 genome, viral RNA was extracted from a known amount of purified virions, as determined by RT activity, using the Qiagen "QIAamp Viral RNA mini kit" as described above. To control for loss of RNA during the extraction, 1 µg of *in vitro* transcribed GAPDH RNA (pos. 61 to 346, see section 2.3.5) was added to the virions prior to extraction. For the HIV-2 specific RT-PCR, 10 µl of neat or diluted viral RNA was mixed with 25 µl of 2 × BioMix Red (Bioline), 1 µl of 20 µM HIV-2 Psi F, 1 µl of 20 µM HIV-2 Psi R, 12 µl of H_2O and 1 µl of AMV RT (Promega). The GAPDH control RT-PCR was performed similarly using primers 5' GGCACCGTCAAGGCTGAG 3' and 5' AATTAACCCTCAC-TAAAGGACTCCACGACGTACTC 3'. A no RT control was set up in parallel. Reactions were reverse-transcribed and amplified in a one-step RT-PCR using the following program: 47 °C for 1 h, 95 °C for 2 min, [95 °C for 30 sec, 45 °C for 30 sec, 72 °C for 30 sec] for 45 cycles and a final extension at 72 °C for 5 min. RT-PCR products were separated by agarose gel electrophoresis in 1 × TBE.

2.7.11 Computer analysis of nucleic acids

RNA secondary structures were predicted using the version 3.2 of the free energy minimisation algorithm *mfold* (Mathews et al., 1999; Zuker, 2003) available at http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/rna-form1.cgi.

Nucleotide alignments were performed using the basic local alignment search tool or BLAST (Altschul et al., 1997) available at http://blast.ncbi.nlm.nih.gov/Blast.cgi.

2.8 Preparation and analysis of proteins

2.8.1 Preparation of proteins from transfected cells

Transfected cells were washed once with ice-cold PbS, trypsinised using 1 × trypsin-EDTA solution and pelleted by centrifugation at 300 × g for 5 min. The cell pellet was resuspended in 100 µl PbS and 100 µl of 2 × protein loading buffer (100 mM Tris-HCl pH 6.8, 200 mM β-mercaptoethanol, 4 % (w:v) SDS, 0.2 % bromophenol blue and 20 % (v:v) glycerol) was added. The cells were sonicated five times for 1 min at 50 % power using a XL2020 sonicator (Heat Systems) prior to analysis by SDS-polyacrylamide gel electrophoresis (PAGE) as described below.

2.8.2 Preparation of proteins from virions

Virions were purified through 20 % sucrose cushion as described in section 2.6.1 and resuspended in 100 µl PbS. 100 µl of 2 × protein loading buffer was added and virions proteins were analysed by SDS-PAGE.

2.8.3 SDS-PAGE

SDS-PAGE was used to analyse proteins levels. Samples were heat-denatured at 95 °C for 3 min prior to loading on 10 to 15 % polyacrylamide gel prepared as described in Molecular Cloning (Sambrook and Russell, 2001). Prestained protein markers (GE-Healthcare) were run in parallel for size determination. When a broad range resolution was required, 4-20 % gradient gels (Pierce) were used according to the manufacturer's instructions.

2.8.4 Western blot analysis

Following separation on a SDS-PAGE, the proteins were transferred onto nitrocellulose membrane (Hybond-C extra, GE Healthcare) in 1 × transfer buffer (25 mM Tris-HCl, 200 mM glycine, 20 % (v:v) MeOH) overnight at 20 Volts or for 2 h at 60 Volts. Membranes were blocked in PbS, 0.05 % Tween-20, 5 % milk for at least 30 min at rt prior to incubation with the primary antibody. Antibodies were diluted as recommended by the manufacturer in the blocking solution. Following incubation at rt for 1 h with gentle agitation, the blot was washed three times for 10 min in PbS, 0.05 % Tween-20 and incubated with the horseradish peroxidase (HRP)-conjugated secondary antibody for 1 h at rt. The blot was washed three times in PbS, 0.05 % Tween-20, developed using the ECL reagents (Pierce) and exposed to X-ray film.

2.8.5 Reverse Transcriptase assay

The reverse transcriptase assay was first described by Potts (Potts, 1990). 10 µl of sample was mixed with 50 µl of RT solution (80 mM Tris-HCl ph 7.8, 120 mM KCl, 3.2 mM DTT, 7.92 mM $MgCl_2$, 80 µg/ml Poly(A), 2.52 µg/ml oligo dT (12-18), 4 % Nonidet-P40, 0.05 µCi [α-^{32}P] dTTP) and incubated at 37 °C for 1 h 30 min. 5 µl were spotted on chromatography paper (DE81 anion exchanger, Whatman), air dried for 15 min and washed four times in 2 × SSC and twice in 96 % EtOH. The paper was dried and radioactivity was counted on a Matrix 96 direct beta counter (Packard).

3 Relationship between genomic RNA dimerisation, packaging and viral infectivity in HIV-2

3.1 Introduction

Retroviruses encapsidate two copies of positive sense single-stranded genomic RNA. Encapsidation is very specific, as the virus has to select and package the full length genomic RNA in preference to the vast excess of cellular and viral subgenomic RNA species. In HIV-1, the Gag structural protein interacts with a highly structured RNA motif located downstream of the splice donor, ensuring specificity of encapsidation for full length RNAs (Aldovini and Young, 1990; Baudin et al., 1993; Berkowitz et al., 1995; Clever et al., 1995; Dannull et al., 1994; Gorelick et al., 1990; Harrison and Lever, 1992; Hayashi et al., 1992; Lever et al., 1989; Luban and Goff, 1994; McBride and Panganiban, 1996).

In the case of HIV-2, the process is less well understood. Although some reports have suggested that sequences downstream of the splice donor are required for RNA packaging (Arya et al., 1998; Garzino-Demo et al., 1995; Poeschla et al., 1998), the main packaging determinant (Psi) was shown to be located upstream of the major splice donor (Griffin et al., 2001; McCann and Lever, 1997). As a result, HIV-2 has been shown to preferentially package its genome in a *cis* rather than a *trans* manner in order to specifically encapsidate the unspliced genomic RNA (Kaye and Lever, 1999). The structural Gag protein is therefore translated from the full length unspliced RNA and encapsidates genomic RNAs present in the same pool of RNAs that were translated.

This *cis* mechanism, combined with a limited availability of the Gag polyprotein, provides specificity (Griffin et al., 2001; Kaye and Lever, 1999). It has also been proposed that the formation of an extended stem-loop 1 (SL-1) structure in the HIV-2 5′ leader, which overlaps with the packaging signal, is required for efficient genome encapsidation and viral replication (Lanchy and Lodmell, 2007), providing evidence that specific RNA structures are implicated in HIV-2 RNA packaging.

Retroviruses are unique amongst other viruses as they have a diploid genome. In HIV-1, genome dimerisation has been shown to occur by loop-loop interaction via a palindromic sequence termed the dimer initiation site (DIS) and subsequent formation of a 'kissing-loop' complex (Clever et al., 1996; Haddrick et al., 1996; Laughrea and Jette, 1994, 1996; Mujeeb et al., 1998; Muriaux et al., 1996b; Paillart et al., 1994; Paillart et al., 1996b; Paillart et al., 1997; Skripkin et al., 1994).

In HIV-2, the sequence initiating dimerisation has not been clearly identified. A palindromic sequence (*pal*) within the encapsidation signal was shown to be important for the regulation of the dimerisation process *in vitro* (Baig et al., 2007; Lanchy et al., 2003a). In addition, several *in vitro* studies proposed that a palindrome located in the loop of SL-1, immediately downstream of Psi, and structurally analogous to the HIV-1 DIS (Laughrea and Jette, 1994; Skripkin et al., 1994), could act as the DIS for HIV-2 (Dirac et al., 2001; Lanchy and Lodmell, 2002). Additional sequences in the PBS were suggested to act as dimerisation signals (Jossinet et al., 2001) but subsequently shown only to promote the formation of a weak dimer in contrast to studies implicating the putative DIS (Lanchy and Lodmell, 2002). To date, the presence of RNA dimers in HIV-2 virions has not been demonstrated.

Since the *cis*-acting RNA elements required for retroviral RNA dimerisation and encapsidation are often located in close proximity in the 5′ RNA leader, it has been

suggested that these two processes may be interdependent (Lanchy et al., 2003a; Paillart et al., 1996a; Russell et al., 2003a; Russell et al., 2004; Sakuragi et al., 2003; Torrent et al., 1994; Whitney and Wainberg, 2006). In addition, the nucleocapsid domain of the Gag polyprotein of several retroviruses has been proposed to act as a *trans*-acting factor involved in both genome encapsidation (Aldovini and Young, 1990; Berkowitz et al., 1995; Fu et al., 2007; Gorelick et al., 1988; Gorelick et al., 1990; Meric and Spahr, 1986; Rein et al., 1994; Schwartz et al., 1997) and dimerisation (Bonnet-Mathoniere et al., 1996; Darlix et al., 1990; Feng et al., 1999; Feng et al., 1996b; Lanchy et al., 2003b; Laughrea et al., 2001; Muriaux et al., 1996a; Prats et al., 1990). In MLV too, dimerisation and packaging are linked since mRNA molecules containing the packaging signal were shown to be encapsidated as dimers (Hibbert et al., 2004).

Interestingly, several *in vitro* studies have shown that long-range interactions can regulate RNA dimerisation in HIV-1 and HIV-2 (Berkhout et al., 2002; Dirac et al., 2002; Huthoff and Berkhout, 2001; Lanchy et al., 2003a) and it has been proposed that RNA packaging may also be regulated through alternative RNA conformations (Lanchy et al., 2003a; Paillart et al., 2002).

However, in HIV-2, a clear relationship between genome dimerisation and encapsidation has not been established and it remains to be demonstrated whether the dimerisation of the genome is a prerequisite for its encapsidation, as proposed for HIV-1 (for review see Russell et al., 2004).

Efficient genome dimerisation and encapsidation have been shown to be important for HIV infectivity. Mutations that impair RNA packaging often result in a delay or defect in replication and a reduced infectivity in HIV-1 (Aldovini and Young, 1990; Clavel and Orenstein, 1990; Clever and Parslow, 1997; Gorelick et al., 1990; Lever et al., 1989). Similarly, genomic RNA dimerisation has been shown to influence viral

infectivity, as HIV-1 mutants with a reduction in genome dimerisation also had a decreased infectivity (Clever and Parslow, 1997; Laughrea et al., 1997; Russell et al., 2003b). Interestingly, a mutation that disrupted the HIV-1 DIS base-pairing did not affect genome dimerisation but resulted in a decrease in the level of RNA packaged in the virions and a delay in viral replication (Berkhout and van Wamel, 1996). While efficient RNA packaging has been shown to be required for HIV-2 replication (Griffin et al., 2001; Lanchy and Lodmell, 2007; McCann and Lever, 1997), it has not been determined whether the same is true for genome dimerisation.

In this study, the relationship between genome dimerisation, packaging and viral infectivity in HIV-2 was investigated. The roles of the Psi and putative DIS palindromes in HIV-2 genome dimerisation and viral infectivity were analysed. Using site-directed mutagenesis, deletions and substitutions were introduced in Psi and the putative DIS palindrome and genomic RNA packaging and dimerisation, protein production and viral infectivity were assessed. It was shown that deletion of Psi rendered the genomic RNA monomeric in virion produced in cell culture (*in virio*) as well as *in vitro*. Furthermore, the results demonstrated that the putative DIS palindrome was not required for genomic RNA dimerisation. Instead, mutations of the Psi palindrome *pal* resulted in a significant reduction in the amount of dimeric genome present in the virions and a decrease in packaging efficiencies. In addition, replication in T-cells was impaired in all dimerisation-deficient mutants. Nonetheless, viral infectivity was found to be affected independently of the reduction in RNA packaging and despite normal Gag protein production and processing.

3.2 Deletion of the main packaging determinant renders the RNA monomeric in cell-culture as well as *in vitro*

The main HIV-2 packaging determinant was identified upstream of the major splice donor (Griffin et al., 2001), at position 380-408 of the HIV-2_{ROD} genome (Figure 3.1A and B). Deletion of this 28 nt sequence (DM mutant, Figure 3.1A and C) was found to greatly reduce genomic RNA packaging (Griffin et al., 2001). Since dimerisation and encapsidation of the genome have been proposed to be linked in HIV-1 and HIV-2, it was decided to analyse the native state of the wild type and DM mutant RNA transcripts as well as genomic RNAs (Figure 3.2). First, the level of dimeric RNA was analysed in an *in vitro* dimerisation assay. This kind of assay had been successfully used to identify the sequences required for HIV-1 genome dimerisation (Laughrea and Jette, 1994; Skripkin et al., 1994) and to predict a putative DIS in HIV-2 (Dirac et al., 2001; Lanchy and Lodmell, 2002). Using the conditions described by Dirac and co-workers (Dirac et al., 2001) and detailed in section 2.7.9, the deletion of the DM sequence was found to reduce the *in vitro* dimerisation of a 130 nt RNA transcript comprising nucleotides 350 to 480 of the HIV-2_{ROD} genome by approximately twofold (Figure 3.2A and B). Similarly, when the genomic RNA from purified virions was analysed by native northern blot (Figure 3.2C), the amount of dimeric genome in the DM mutant was significantly decreased from 80 % (wild type) to 30 % (Figure 3.2D).

This was the first evidence of dimeric genomes in HIV-2 virions and confirmed that HIV-2 genomic RNA, like that of all exogenous retroviruses studied to date, is dimeric within viral particles. In addition, these results implied that the HIV-2 packaging signal may contain sequences that are important for RNA dimerisation.

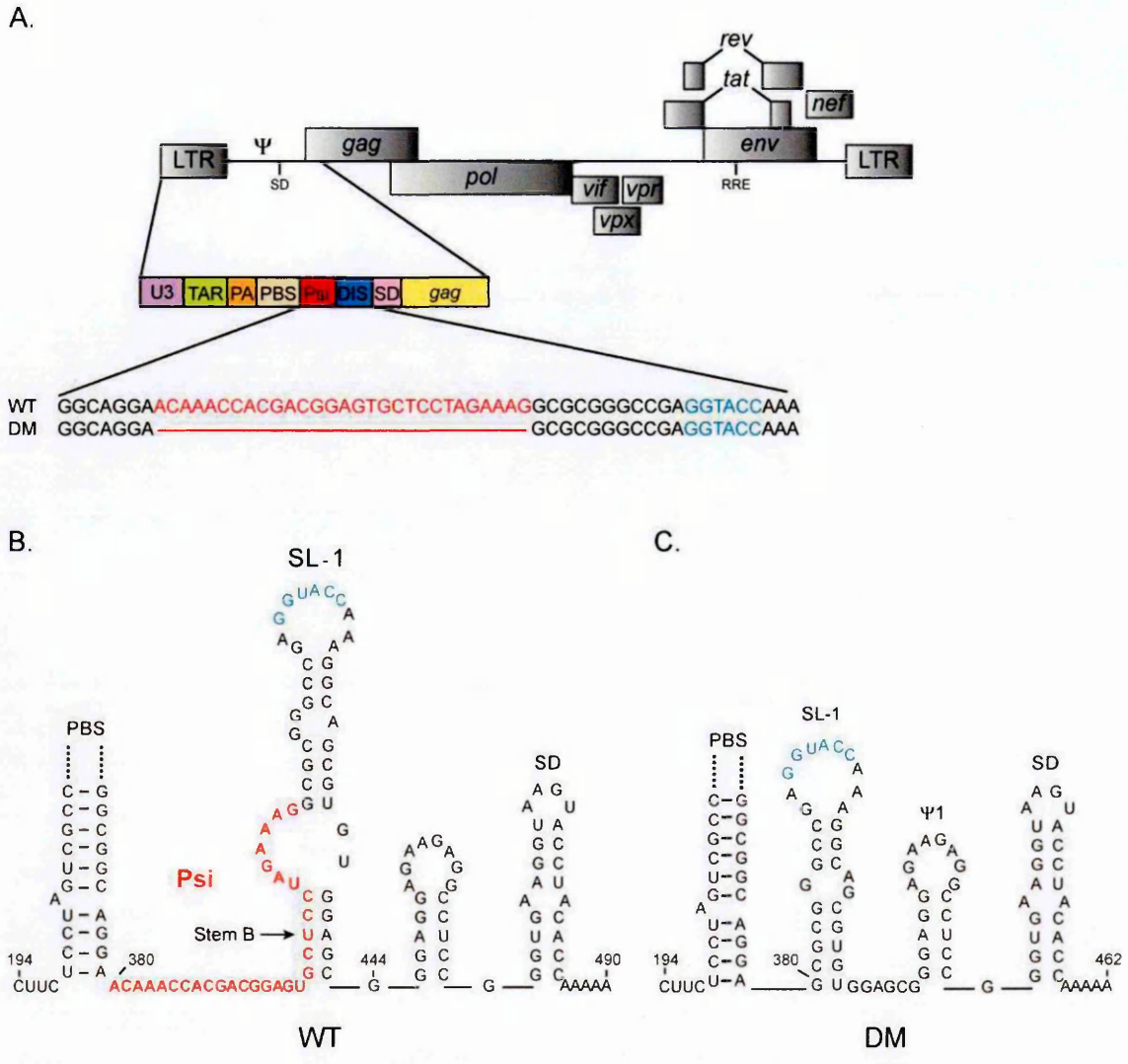


Figure 3.1: Genomic and structural context of the HIV-2 Psi deletion mutant.
(A) Schematic genetic organisation of the HIV-2_{ROD} molecular clone, pSVR. LTR, long terminal repeat; Ψ, packaging signal (Psi); SD, splice donor; RRE, Rev responsive element; U3, unique 3′; R, repeat; TAR, *Trans*-activation responsive; PA, Poly A; PBS, Primer binding site; DIS, Putative dimer initiation site; WT, wild type; DM, DM deletion mutant. The packaging signal and putative DIS are indicated in red and blue respectively. (B) Structure of the WT HIV-2 5′ RNA leader (from position 370 to 490) based on biochemical analysis (Baig et al., 2007; Damgaard et al., 1998; Dirac et al., 2001; Lanchy et al., 2004) and minimal free-energy modelling using *mfold* (Mathews et al., 1999; Zuker, 2003). The packaging signal Psi is indicated in red and the putative DIS in blue. (C) Structure of the DM mutant 5′ RNA leader (position 370 to 462) predicted by minimal free-energy modelling using *mfold*. The putative DIS is in blue.

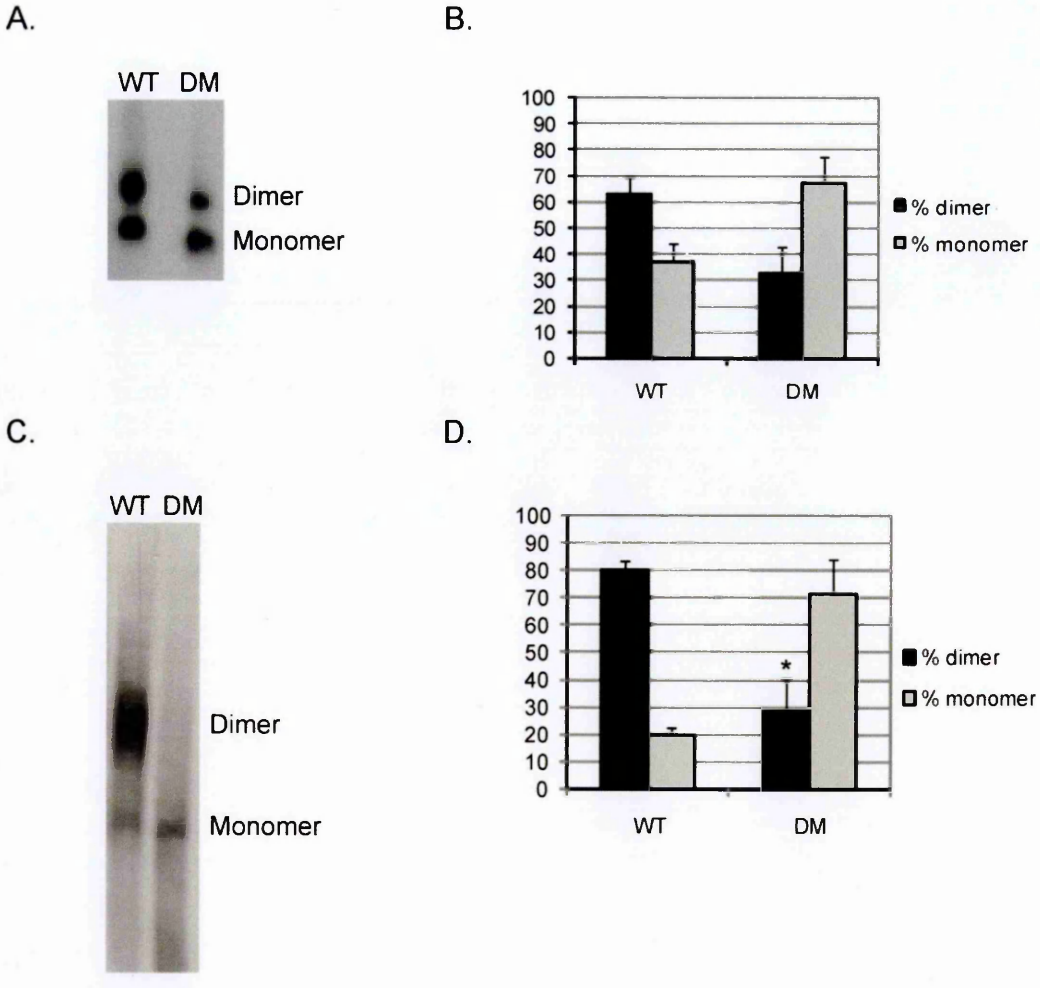


Figure 3.2: Deletion of the HIV-2 packaging signal significantly reduces genomic RNA dimerisation.

(A) Representative *in vitro* dimerisation assay using wild type (WT) and DM mutant (DM) RNA transcripts (position 350-480 of the HIV-2_{ROD} genome). The positions of the dimeric and monomeric RNA species are indicated. (B) Quantification of monomeric and dimeric RNA species for the WT and DM transcripts using the ImageJ software. Average and standard deviation (sd) from three independent experiments are shown. (C) Native northern blot analysis of WT and DM genomic RNA. 2000 cpm of viral RNA extracted from purified virions and normalised on RT activity was analysed by native northern blot using the biotinylated KS2ΨKE riboprobe. The positions of the RNA dimers and monomers are indicated. (D) Quantification of the monomeric and dimeric RNA genomes for the WT and DM viruses using ImageJ. Average and sd of three independent experiments are shown. *, two-tailed unpaired Student t test p value < 0.05.

Interestingly, the reduction in RNA dimerisation was greater when analysed in the context of the whole genome, where factors such as long-range interactions and the viral NC protein are present which have been shown to influence genome dimerisation in several retroviruses (Abbink et al., 2005; Clever et al., 1995; Darlix et al., 1990; Dirac et al., 2002; Lanchy et al., 2003b; Prats et al., 1990).

3.3 The Psi palindrome *pal*, but not the putative DIS palindrome, is required for RNA dimerisation *in virio*

Although we had established that deletion of the HIV-2 encapsidation determinant affected genome dimerisation, the exact location of the *cis*-acting element required for genomic RNA dimerisation remained unknown. Whilst the Psi sequence was partly unstructured, it also extended into the SL-1 structure which contained the putative DIS (Figure 3.1B). Hence, deletion of the Psi sequence in the DM mutant was predicted to result in an alteration of the SL-1 structure which might prevent a loop-loop interaction through occluding the putative DIS palindromic motif and because of its closer proximity to the PBS stem-loop structure (Figure 3.1C). The predicted structure of the dimerisation-deficient DM mutant did however leave intact the putative dimerisation site in the loop of SL-1, suggesting that the putative DIS palindrome may not be critical for genome dimerisation *in virio*. The DM region also contained a palindrome (*pal*) that had been shown to promote RNA dimerisation *in vitro* (Lanchy et al., 2003a).

Hence, in order to determine which palindromic sequence was necessary and sufficient to promote RNA dimerisation in HIV-2, both palindromic sequences were mutated by site-directed mutagenesis (Figure 3.3A). The mutant Pal carried a deletion of the 10 nt Psi palindrome *pal* (Figure 3.3A, green residues). In SM1, the first three bases of the

terminal palindrome (putative DIS) were substituted so that it was no longer a self-complementary sequence (Figure 3.3A, blue residues, yellow highlight). SM2 was a similar substitution of the first four bases of *pal* (Figure 3.3A, green residues, yellow highlight). SM3 corresponded to a double mutation combining SM1 and SM2 (Figure 3.3A, green and blue residues, yellow highlights).

The minimal free-energy structures of the 5′ RNA leader of these mutants were predicted using the algorithm *mfold* (Figure 3.3B-D). Deletion of *pal* was predicted to diminish the length of SL-1 and prevent the formation of the bulge in SL-1 (Figure 3.3B). However, unlike in the DM mutant (Figure 3.1C), the distance between the PBS stem-loop and SL-1 was predicted to be maintained by the presence of a single-stranded region (Figure 3.3B, red). The SM1 mutation was not predicted to affect the structure since the mutated residues were located in the loop of SL-1 (Figure 3.3C). Substitution of the residues 392-395 in SM2 altered the conformation of SL-1 with the formation of an extended SL-1 and the disappearance of the neighbouring Ψ1 stem-loop (Figure 3.3D). The formation of stem B, which was proposed to be necessary for viral replication (Lanchy and Lodmell, 2007), was however maintained.

First, the impact of these mutations on *in vitro* RNA dimerisation was analysed using the *in vitro* assay used for the DM mutant and described in section 2.7.9 (Figure 3.4A and B). The wild type RNA transcript was approximately 60 % dimeric in this assay. Deletion or substitution of *pal* in the mutants Pal and SM2 reduced the proportion of dimeric RNA by approximately twofold (Figure 3.4B), but this reduction was not statistically significant as assessed by a Student t test and assuming a normal distribution of the data. More striking however was the effect of the SM1 mutation, which completely abrogated dimerisation of the RNA *in vitro* (Figure 3.4A and B). This phenotype appeared dominant as the SM3 double mutant displayed a phenotype similar

to that of SM1. The dimerisation of the genomic RNA in virions was then assessed in order to confirm whether the putative DIS was indeed the main requirement for HIV-2 dimerisation (Figure 3.4C and D). As for the DM mutant, viral RNA extracted from purified virion was analysed by native northern blot where RNA input was normalised on the RT activity of the purified virions (Figure 3.4C). Whilst the wild type genomic RNA appeared 80 % dimeric within the virions, a significant decrease in the level of dimeric genomes was observed when the palindrome *pal* was mutated in the Pal and SM2 mutants (Figure 3.4D), suggesting that *pal* plays a role in genomic RNA dimerisation. Surprisingly, in contrast with the data obtained *in vitro*, the SM1 mutation had no effect on genome dimerisation and the SM3 double mutant had a phenotype similar to that of the SM2 mutant (Figure 3.4C and D). These results suggested that even though the putative DIS can promote the *in vitro* dimerisation of a short RNA transcript, in the context of the whole viral genome, it is not required to initiate the dimerisation process. Interestingly, northern blot analysis of the genomic RNA from HIV-1 DIS mutants also revealed that the HIV-1 DIS was not essential for genome dimerisation (Berkhout and van Wamel, 1996), whilst *in vitro* studies proposed the DIS as being the primary dimerisation site for HIV-1 (Laughrea and Jette, 1994; Skripkin et al., 1994).

These discrepancies between the *in vitro* and the cell culture-based assays may be due in part to the different structure adopted by the RNA in the context of a 130 nt transcript as compared to the full-length genomic RNA. Long-range interactions in the HIV-1 and HIV-2 5′ RNA leader have been demonstrated to influence the dimerisation process, at least *in vitro* (Abbink et al., 2005; Dirac et al., 2002; Lanchy et al., 2003b), and the size of the RNA transcripts used has been shown to greatly influence the outcome of *in vitro* dimerisation assays (Lanchy and Lodmell, 2002).

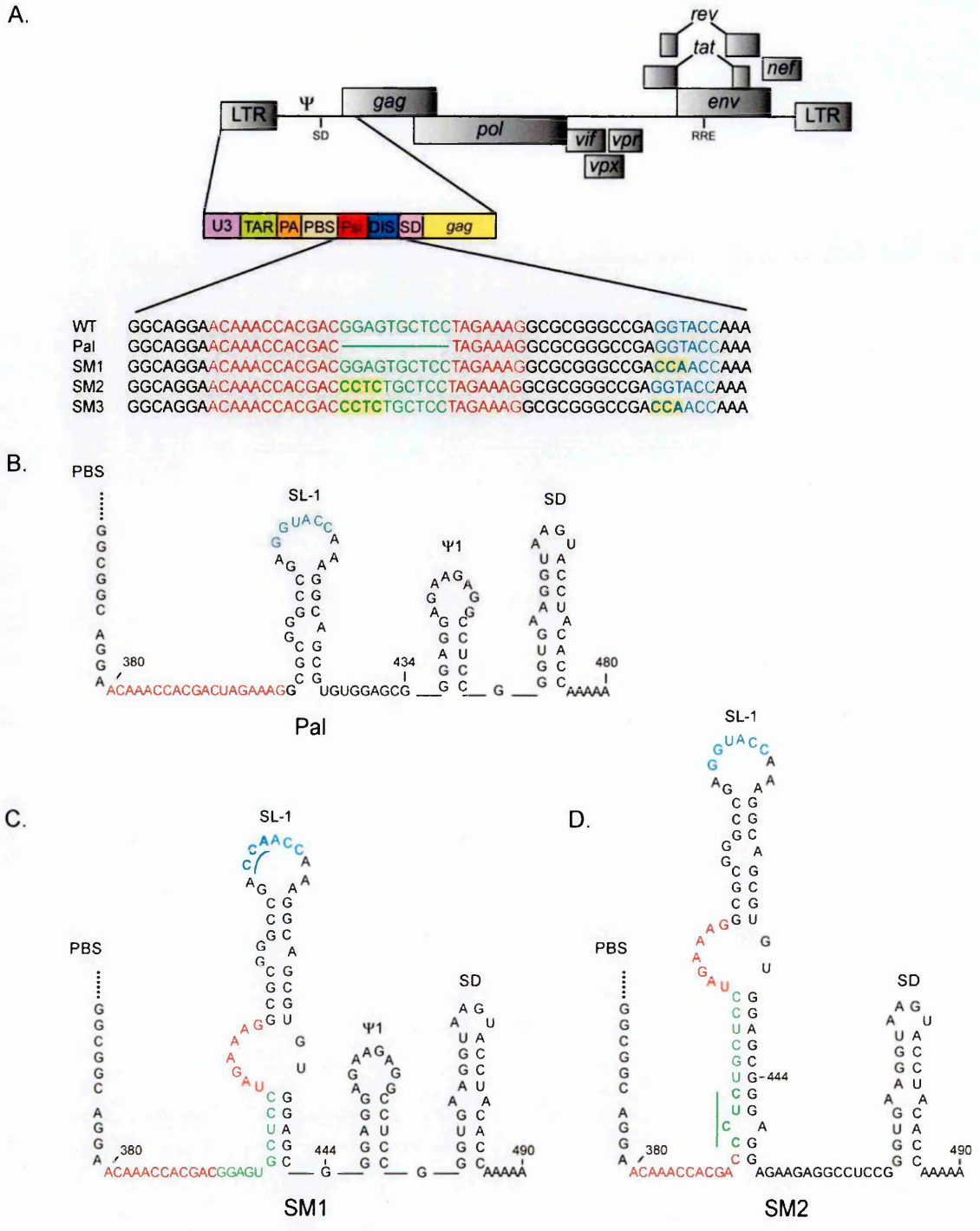


Figure 3.3: Genomic and structural context of the Psi and putative DIS palindrome mutations.

(A) Genomic context of the mutations of the *pal* and putative DIS palindromes. LTR, long terminal repeat; Ψ, packaging signal (Psi); SD, splice donor; RRE, Rev responsive element; U3, unique 3′; R, repeat; TAR, *Trans*-activation responsive; PA, PolyA; PBS, Primer binding site; DIS, Putative dimer initiation site; WT, wild type. (B-D) Predicted structures of the 5′ RNA leader (position 370 to 490) for mutants Pal (B), SM1 (C) and SM2 (D) using *mfold*. The Psi region, Psi palindrome *pal* and putative DIS are indicated in red, green and blue respectively. The mutated residues are highlighted in yellow (A) or underlined (C-D).

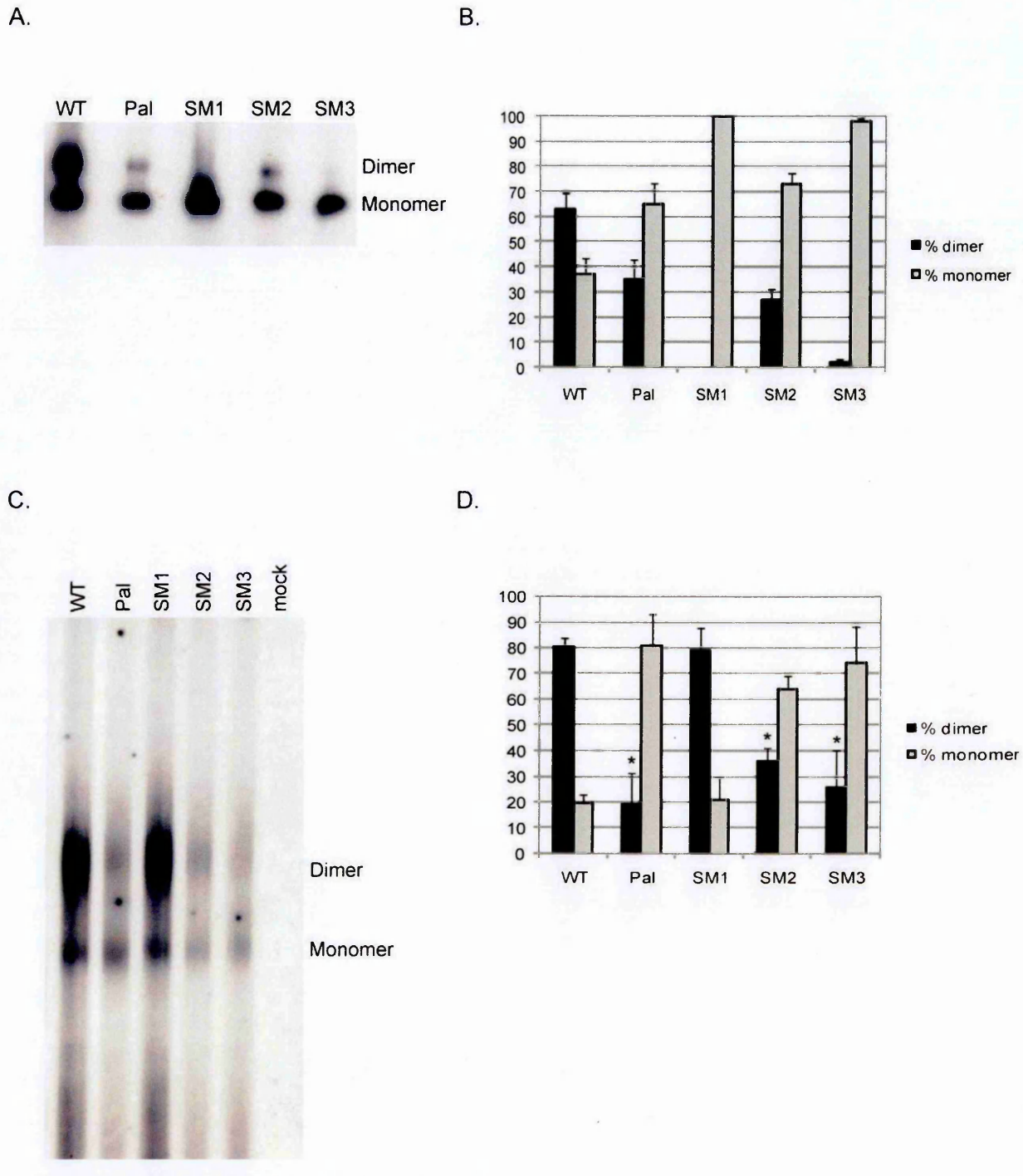


Figure 3.4: The Psi palindrome *pal*, not the putative DIS palindrome, is required for dimerisation of the HIV-2 genomic RNA.

(A) *In vitro* dimerisation assay using a 130 nt RNA transcript (pos. 350-480 of HIV-2_{ROD}). The positions of the monomer and dimer bands are indicated. (B) Quantitative analysis of the *in vitro* dimerisation assay using the ImageJ software. Average and sd of three independent experiments are plotted. (C) Native northern blot analysis of viral RNA extracted from purified virion, using 2000 cpm of viral RNA as normalised on RT activity. The positions of the monomer and dimer RNA species are indicated. (D) Quantification of the dimer and monomer RNA species in the northern blot using ImageJ. Average and sd of three independent experiments are represented. *, two-tailed unpaired Student t test p value < 0.05. (A-D) WT, wild type; mock, mock transfected cells.

In addition, the presence of the NC domain of the Gag polyprotein was proposed to affect dimer formation in HIV-1 (Darlix et al., 1990; Feng et al., 1996b; Fu et al., 1994; Muriaux et al., 1996a) and HIV-2 (Dirac et al., 2002; Lanchy et al., 2003b).

Even though a role for the putative DIS palindrome in the stabilisation or maturation of the dimer cannot be excluded, as the stability of the dimers in the mutant viruses has not been assessed, the palindrome *pal* appeared to represent the critical sequence for HIV-2 genome dimerisation.

3.4 The putative DIS palindrome is dispensable for replication in T-cells

In HIV-1, the DIS has been shown to mediate dimerisation (Laughrea and Jette, 1994; Skripkin et al., 1994), yet whether it is required for viral replication remains unclear (Berkhout and van Wamel, 1996; Hill et al., 2003; Jones et al., 2008; Laughrea et al., 1997). The HIV-1 DIS was shown to be dispensable for replication in PBMCs (Hill et al., 2003) but was required in SUP-T1 T-cells (Berkhout and van Wamel, 1996; Hill et al., 2003), suggesting a cell type-dependent mechanism. Deletions and substitutions of the HIV-1 DIS also reduced the infectious titres by at least twofold in two T-cell lines, SUP-T1 (Paillart et al., 1996a) and MT-4 cells (Laughrea et al., 1997).

Since the current study showed that the Psi palindrome *pal*, not the putative DIS palindrome, was necessary for HIV-2 genome dimerisation and considering that mutations affecting HIV-2 genome encapsidation also affect viral replication (Griffin et al., 2001; Lanchy and Lodmell, 2007; McCann and Lever, 1997), it was important to measure the viral infectivity of the mutant viruses (Figure 3.5).

A single-round infectivity assay was performed using a reporter cell line, termed GHOST cells. This clonal derivative of HUT78 cells is a human osteosarcoma cell line, constitutively expressing CD4, CXCR4 and CCR5. It contains the green fluorescent protein (*gfp*) gene, the expression of which is controlled by an HIV-2 promoter (Trkola et al., 1998). Upon infection by HIV-2, the Tat protein *trans*-activates the expression of GFP, which can be detected by FACS.

Wild type and mutant viruses were first produced by transfection of Cos-1 cells and harvested at 48 h post-transfection. An increasing amount of virus, as determined by an RT assay, was added to the GHOST cells. At 72 h post-infection, the percentage of GFP-expressing cells, which represents the percentage of infected cells, was determined by FACS (Figure 3.5A). For safety reason, VSV-G pseudotyped viruses were used in this experiment. VSV-G pseudotyped wild type HIV-2 and SM1 viruses appeared fully infectious, with up to 90 % of infected cells at the highest input tested. However, viruses containing mutations within the Psi region (DM, Pal, SM2 and SM3) all failed to efficiently infect the reporter cells, with a maximum of 30 % to 50 % of cells being infected. A reversion of the DM mutation was introduced into the pSVRΔNBDM molecular clone and the infectivity of the resulting virus (DM rev) was determined. The DM rev virus was able to infect cells as efficiently as wild type virus (Figure 3.5A) demonstrating that the infectivity defect observed for the DM mutant was the sole result of the DM deletion and not due to the introduction of another mutation in the proviral backbone during the cloning process. As a control, each individual plasmid was transfected alone. Neither the envelope-deleted proviral construct, nor the VSV-G plasmid, triggered expression of GFP in the reporter cells.

The assay was repeated once using a different cell line to produce the viruses, C33-A, which is a human carcinoma cell line (Figure 3.5B), since it had been shown that for

certain viruses, such as hCMV, the producer cell line could affect the cellular response to infection by influencing the mode of entry of the virus (Wang et al., 2007). Although there was no evidence of such a producer cell type-dependent mode of entry for HIV, it was decided to test an alternative cell line.

A similar pattern was observed when viruses were produced in C33-A or Cos-1 cells, even though the percentage of infected cells was lower when C33-A cells were used (compare Figure 3.5A and B). This difference was likely due to the difference in number of passage of the GHOST cells and the level of expression of the CD4 receptor and CCR5 and CXCR4 co-receptors as it was observed that the percentage of infected cells was higher when the cells had been grown for at least five weeks in selection media prior to use.

The results obtained in C33-A cells showed that the decrease in infectivity of the Psi mutant viruses observed using Cos-1 cells represented a genuine defect of the mutated virus.

To confirm these findings in a system that allows the virus to replicate, T-cells were infected with equal amounts of wild type and mutant viruses and the replication kinetics of the viruses were established (Figure 3.6). Viruses were produced by transfection of Cos-1 cells with the appropriate proviral clones and normalised on RT activity prior to the infection of Jurkat T-cells (Figure 3.6A). The replication was followed by measuring the RT activity in the culture supernatant every 4 days, when the cells were passaged one in two. Following an initial peak of infection at 15 days post-infection, the wild type virus continued to replicate in the Jurkat cells up to 50 days post-infection. The replication of the SM1 mutant was comparable to that of the wild type, with an initial peak of infection at day 15 and a second one at day 50 post-infection, demonstrating that the SL-1 terminal palindrome was dispensable for HIV-2 replication in T-cells.

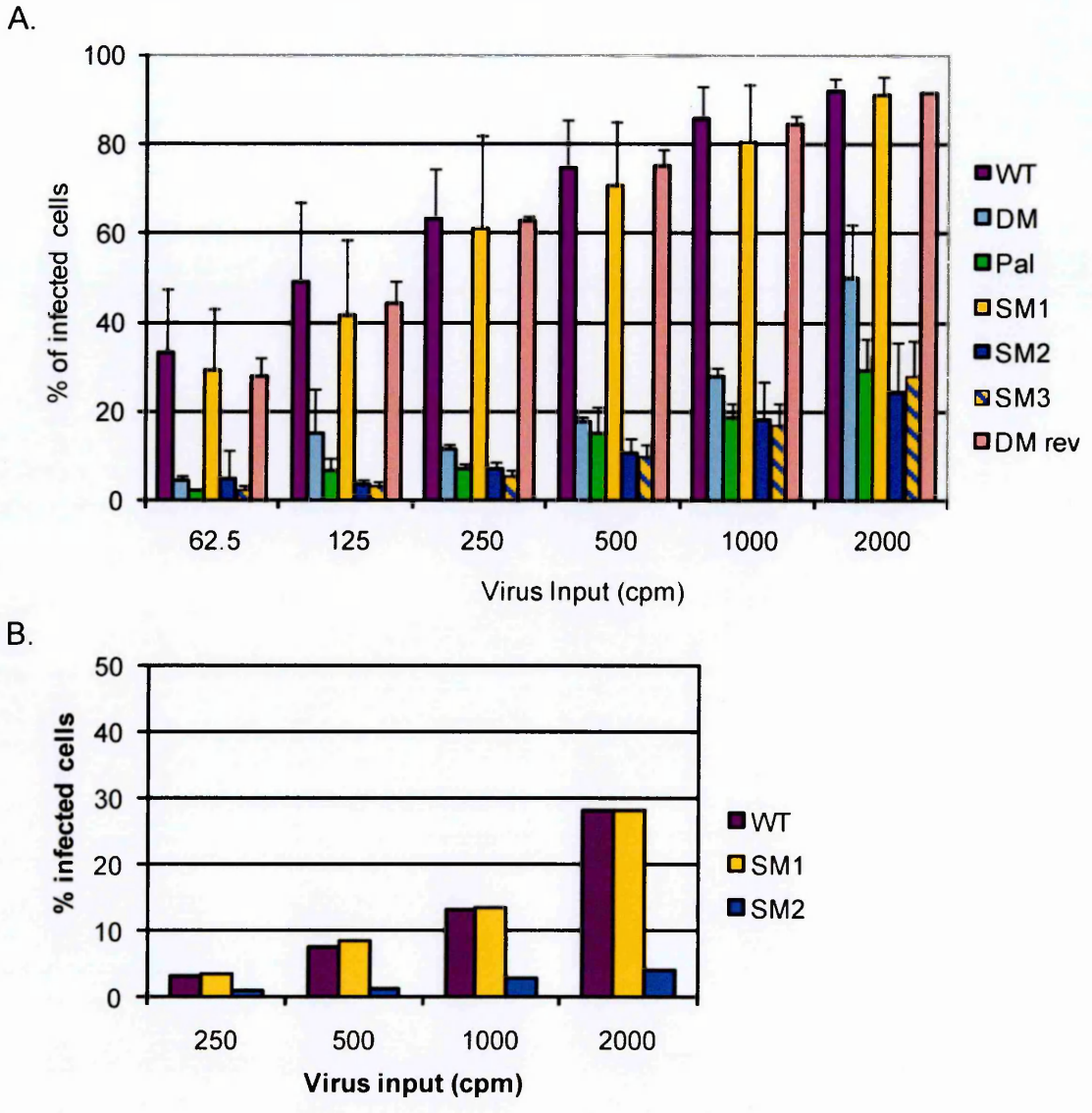


Figure 3.5: Infectivity of wild type and mutant HIV-2 produced in Cos-1 and C33-A cells.

Single round infectivity assay using the GHOST reporter cell line, a $CD4^+$ $CXCR4^+$ $CCR5^+$ human osteosarcoma cell line expressing the *gfp* gene under the control of the HIV-2 promoter. (A) VSV-G pseudotyped viruses were produced by co-transfection of 1.5×10^6 Cos-1 cells with 5 µg of envelope-deleted proviral construct and 1 µg pCMV-VSVG (DEAE-dextran method). Virions were purified through sucrose cushion and normalised on RT activity. 2×10^5 GHOST cells were infected with increasing amounts of virus. At 72 h post-infection, the percentage of GFP-expressing cells, representing the proportion of infected cells, was determined by FACS. WT, wild type; DM rev, cloned reversion of the DM mutation. Average and sd of three independent experiments is plotted. (B) Full-length viruses were produced by transfection of 1.5×10^6 C33-A cells with 30 µg of proviral construct using the calcium-phosphate transfection method. GHOST cells were infected as described above. Data from a single experiment are shown.

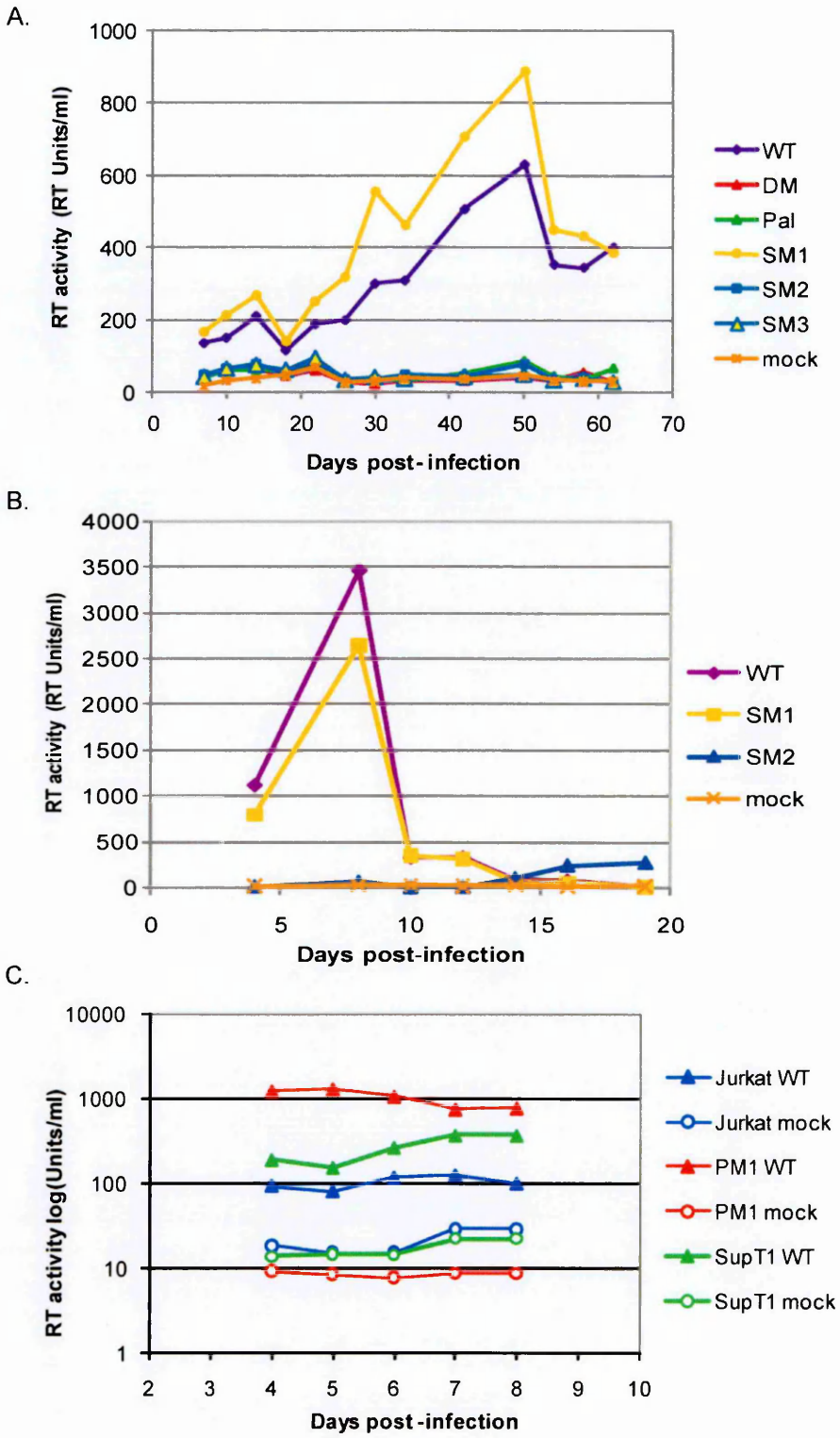


Figure 3.6: Replication of wild type and mutant HIV-2 in T-cells.

1×10^6 Jurkat (A) and 1×10^6 PM1 (B) cells were respectively infected with 10000 cpm and 2500 cpm of purified virions produced by Cos-1 cells transfection. (C) 1×10^6 Jurkat, PM1 and SUP-T1 cells were infected with 7000 cpm of purified wild type HIV-2 produced by transfection of Cos-1 cells. (A-C) Viral replication was followed by measuring the RT activity in a 10 µl aliquots of culture supernatant in duplicate. WT, wild type; mock, mock transfected cells. Data from a single (B and C) or two (A) independent experiments are shown.

In contrast, deletion of the packaging signal (DM), deletion (Pal) and substitution (SM2) of *pal* and double mutation of the putative DIS and *pal* (SM3) all impaired viral replication of HIV-2.

The replication of two mutant viruses and the wild type was also measured in PM1 T-cells (Figure 3.6B) as it was observed that these cells could support HIV-2 replication more effectively than Jurkat or SUP-T1 cells (Figure 3.6C). Similar results were obtained in PM1 cells compared to Jurkat cells, with the wild type and SM1 viruses being able to replicate while the SM2 mutant could not.

Taken together, these data suggested that viruses bearing mutations in the Psi region are non-infectious. Interestingly, the SM2 substitution of the GGAG motif at position 392-395 resulted in the same phenotype as the 28 nt deletion in the DM mutant, suggesting that this purine motif plays an important role in HIV-2 replication.

3.5 Correlation between HIV-2 RNA packaging, dimerisation and replication

Since the mutations were introduced in the 5′ RNA leader, a highly structured region involved in numerous stages of the HIV-2 life cycle, several hypotheses could explain the replication defect observed with the Psi mutants.

The translation of the Gag polyprotein might for example be affected. In their initial analysis, Griffin and colleagues analysed the protein production of the DM mutant by metabolic labelling and found no significant difference with the wild type virus (Griffin et al., 2001). To verify that the other mutations introduced did not affect protein synthesis, viral proteins were harvested from Cos-1 transfected cells and purified virions at 48 h post-transfection and protein synthesis was analysed by western blot using a

panel of HIV-2 positive patient sera (Figure 3.7). No differences in the level of Env proteins (gp125 and gp36) could be detected at 48 h post-transfection between the wild type and mutant viruses. The level of the Gag (p57) polyprotein and cleavage products MA-CA-p2 (≈p41) and CA (p26) also appeared mostly similar between the wild type and mutant viruses. Some variations were observed between the RT (p66) and Gag (p57) levels in the wild type and SM1 compared to the rest of the mutants, but this result was not consistent amongst repeat experiments. It is interesting to note that two of the three isoforms of the Gag polyprotein proposed to be synthesised, namely p57 and p50 (Herbreteau et al., 2005), could be detected in the virion fraction. The ratio of these isoforms was also maintained between the wild type and mutant viruses.

Since the dimerisation-deficient mutants (DM, Pal, SM2 and SM3) failed to replicate in T-cells in spite of a normal protein production, it was of interest to determine whether their packaging efficiencies had been affected by the mutations. Indeed, the inability to replicate in T-cells could result from either a lack of genomic RNA in the incoming virion or the absence of stable dimers in those virions. An RNase protection assay (RPA) was performed to assess the level of HIV-2 genomic and spliced RNA present in the cytoplasm of Cos-1 cells and released virions at 48 h post-transfection. The sizes of unspliced RNA fragments protected by the KS2ΨKE riboprobe are indicated in Figure 3.8A. DNase treated cytoplasmic and virion RNA samples were normalised on the concentration of total cytoplasmic RNA and RT activity of the virus preparation respectively. A representative RPA probing with KS2ΨKE is shown in Figure 3.8B. A GAPDH riboprobe was used to control the loading of the cytoplasmic RNA samples and plasmid DNA contamination was detected using a vector backbone-specific riboprobe (KS2ΨEP).

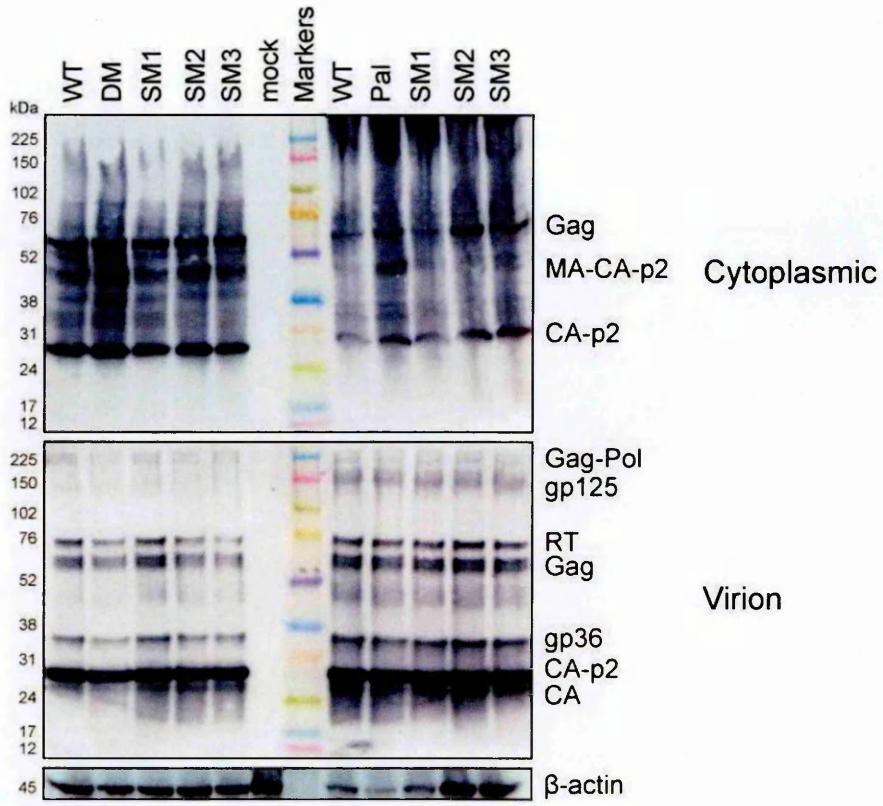


Figure 3.7: Protein expression of WT and mutant viruses.

Cos-1 cells were transfected with WT and mutant proviral constructs and cells and virions were harvested at 48 h post-transfection. 15 µl of the cytoplasmic fraction (top panel) and 500 cpm of the virion fraction (middle panel) were analysed by SDS-PAGE using a 4-20 % gradient gel (Pierce). Protein synthesis was analysed using a 1:1000 dilution of a panel of HIV-2 positive sera (ARP501, NIBSC, UK). Loading was controlled for by probing for β-actin (bottom panel) using a 1:5000 dilution of a mouse monoclonal anti-β-actin (Abcam). 1:1000 dilutions of rabbit anti-human-HRP or goat anti-mouse-HRP conjugated secondary antibodies (Dako) were used. WT, wild type; mock, mock transfected cells. The sizes of the protein markers (full-range Rainbow; GE Healthcare) are indicated on the left. The positions of the HIV-2 proteins are indicated on the right: Gag-Pol (p160), Env SU (gp125) and TM (gp36), RT (p66), Gag (p57), MA-CA-p2 (≈p41), CA-p2 (≈p26) and CA (p26). The multiple bands seen for Gag are thought to represent the different isoforms of the polyprotein (p57, p50 and p44) depending on the initiation codon usage, as described by Herbreteau et al. (2005).

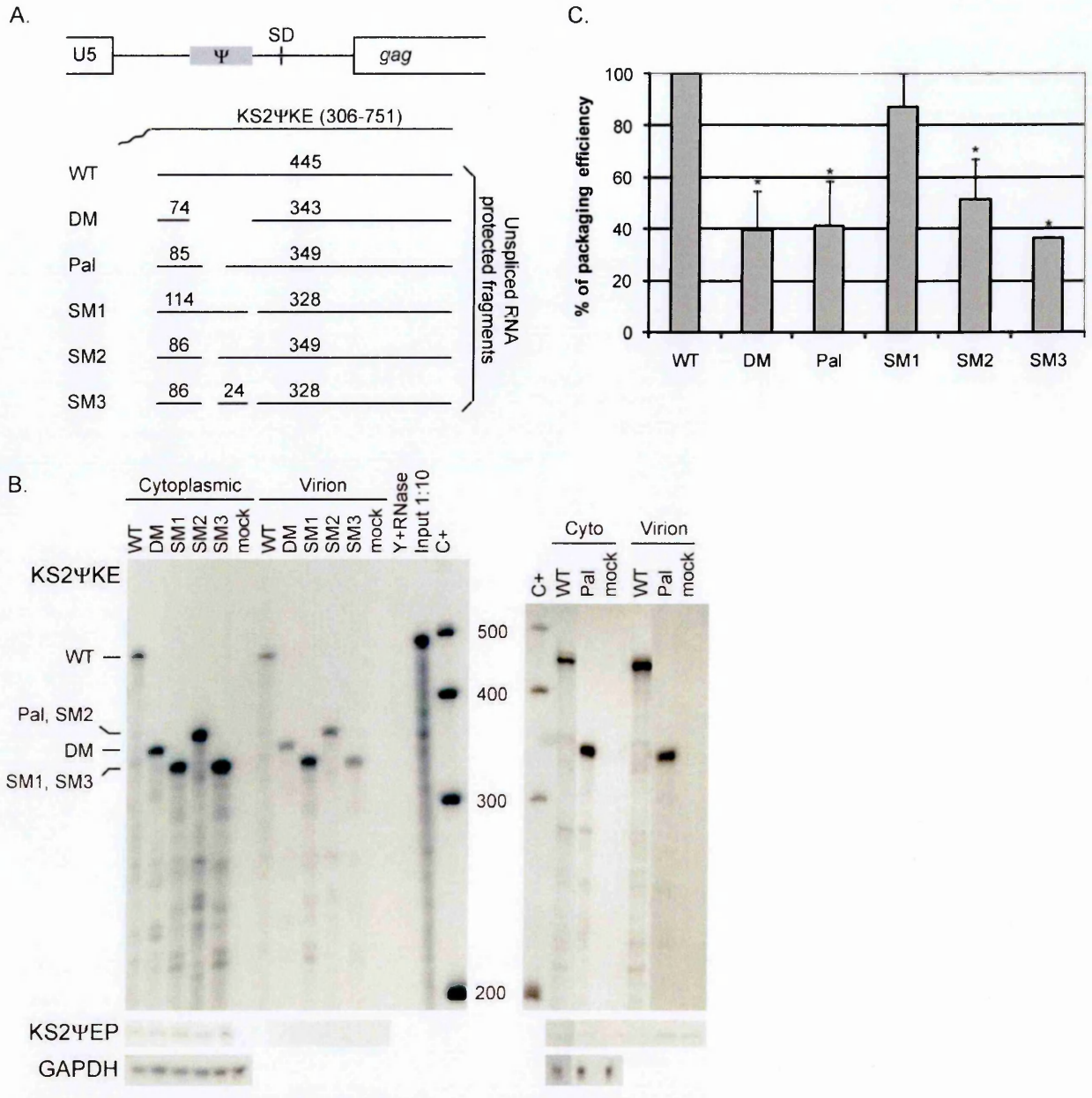


Figure 3.8: Packaging efficiencies of the HIV-2 mutants.

(A) Size of the protected fragments for wild type and mutant genomic RNA using the KS2ΨKE riboprobe. (B) Representative RPAs using the KS2ΨKE riboprobe. 0.5 µg of cytoplasmic RNA and 2500 cpm virion RNA were used with 2×10^5 cpm of ^{32}P-labelled riboprobe. WT, wild type; mock, mock transfected cells; Y+, yeast RNA + RNase; Input 1:10, 1:10 dilution of yeast RNA - RNAse; C+, Century Plus RNA markers (Ambion). The positions of the main protected fragments are indicated on the left. Detection of plasmid DNA with KS2ΨEP (413 nt protected fragment) and GAPDH internal control (285 nt protected fragment) are shown below. (C) Relative packaging efficiency of the mutants, taken as the ratio of virion RNA to cytoplasmic RNA, relative to that of the wild type. Average and sd of at least two independent experiments is shown. *, two-tailed unpaired Student t test p value < 0.05.

The position of wild type and mutant genomic RNA protected fragments are indicated on the side of the autoradiograph. The packaging efficiency of each virus relative to wild type was calculated and is shown in Figure 3.8C. In accordance with previous findings, deletion of the DM sequence resulted in a significant, yet not as marked, decrease in packaging efficiency compared to wild type. In their initial analysis, Griffin and co-workers described a packaging efficiency of 5.7 ± 1.6 % for the DM mutant (Griffin et al., 2001), whereas in this study, the figure was 40 ± 15 %. The decrease in the relative packaging efficiency was still significant with a Student t test p value less than 0.05. Mutations of *pal* also resulted in a significant decrease of the packaging efficiencies of these viruses, with two- to threefold reduction for Pal and SM2. The SM1 mutation of the putative DIS had however no effect on the packaging efficiency of the virus. The SM3 double mutant displayed a phenotype similar to that of SM2 with a two- to threefold reduction in encapsidation compared to the wild type virus. These results showed a correlation between the ability of the virus to form a dimeric genome, its ability to encapsidate that genome and its replicative fitness.

3.6 Viral infectivity is independent of the amount of genomic RNA packaged

A correlation was observed between the dimerisation, encapsidation and infectivity defect, suggesting that these three processes are closely linked. However, the exact nature of their relationship remained unknown. In particular, it was unclear whether the encapsidation of the genome required dimerisation and if the reduced infectivity resulted from a decrease in RNA packaging, dimerisation or both.

If the diminution in encapsidation, *i.e.* a lack of genomic RNA, was the sole factor responsible for the reduction in viral infectivity, the targeting of an equivalent amount of genomes onto the cells would elicit a similar infectivity for the packaging mutant and wild type virus. The DM virus showed a two- to threefold decrease in RNA packaging relative to the wild type virus (Figure 3.8). Hence, a DM virus input three times larger than that of the wild type would be expected to contain as many copies of the genomic RNA as the wild type virus input. Following this assumption, a single-round infectivity assay was performed using three times more DM virus than wild type virus (Figure 3.9A). The virus inputs were normalised on the RT activity of the virus preparation, an equivalent of 62.5 and 250 cpm being used for the wild type virus and 187.5 and 750 cpm for the DM virus. Semi-quantitative RT-PCR was used to estimate the genomic RNA level for each input tested. A control with no reverse transcriptase was performed to ensure that no plasmid DNA contaminated the RNA samples. For each input tested, the genomic RNA content was equivalent for the wild type and DM virus. Yet, the DM virus still showed a statistically significant diminution in infectivity suggesting that factors other than RNA packaging were affecting the infectivity of the DM mutant.

These results were corroborated by the replication kinetic of the viruses in T-cells. In this experiment, the input of wild type virus was reduced by threefold to approximately 3000 cpm, as determined by an RT assay, while the DM virus input was 9000 cpm. This approach was necessary due to the limited amount of DM virus produced by the transfection of Cos-1 cells. A normal input of wild type virus (9000 cpm) was used in parallel. Jurkat cells were infected with wild type and DM virus and viral replication was followed for three weeks by measuring the RT activity in the culture supernatant (Figure 3.9B).

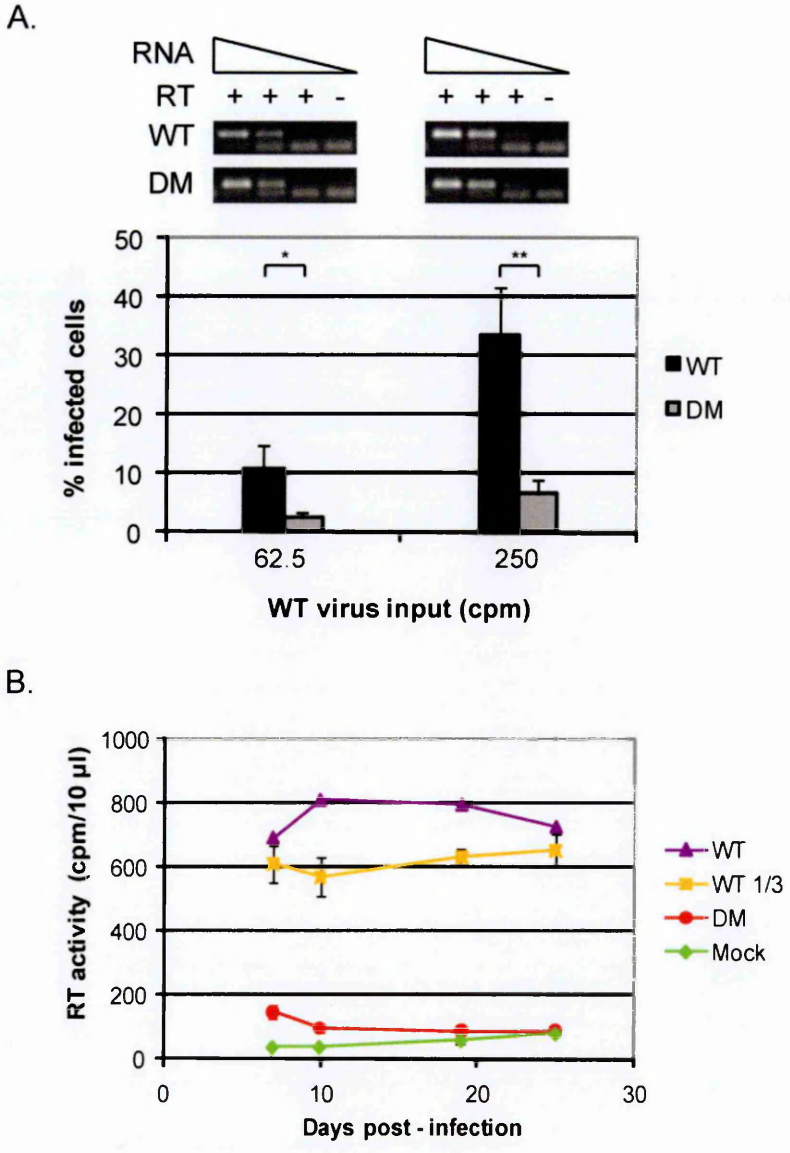


Figure 3.9: The infectivity defect of the DM virus is independent of its packaging defect.

(A) Pseudotyped viruses were produced by co-transfection of 1.5×10^6 Cos-1 cells with 5 µg of pSVRΔNB (WT) or pSVRΔNBDM (DM) and 1 µg pCMV-VSVG, purified through sucrose cushion and normalised on RT activity. 2×10^5 GHOST cells were infected with 62.5 and 250 cpm of WT virus and 187.5 and 750 cpm or DM virus. At 72 h post-infection, the percentage of GFP-positive cells was determined by FACS. In the meantime, the viral RNA from an equivalent fraction of virions as used to infect the cells was extracted for RT-PCR analysis. 10 µl of undiluted, 1:10 and 1:100 dilution of viral RNA were reverse transcribed and amplified using HIV-2 specific primers as described in section 2.7.10. A control reaction without RT was carried out in parallel. Statistical significance was assessed by a two-tailed unpaired Student t test. *, p value < 0.05, **, p value < 0.01. (B) 1.5×10^6 Jurkat cells were infected with either 9000 cpm of purified WT and DM viruses or 3000 cpm of purified WT virus (WT 1/3), as measured by RT activity. At day 7, 10, 19 and 25, two 10 µl aliquots were removed to measure RT activity in duplicate. Mock, mock transfected cells. Average and sd from two independent experiments are shown.

As previously observed, the DM virus failed to replicate in T-cells. Diminishing the input of wild type virus, so that it would have an amount of targeting genome equivalent to that of the DM sample, did not significantly affect the ability of the virus to replicate in Jurkat cells.

3.7 Discussion

The relationship between genome dimerisation, encapsidation and viral replication was investigated. Using native northern blot analysis of the HIV-2 full-length genomic RNA, it was shown that sequences within the HIV-2 packaging signal were involved in the dimerisation of the genome. Deletions of Psi and deletion or substitution in the palindrome *pal* resulted in a significant reduction in the amount of dimeric genomes present in the virions, suggesting that *pal* is the main requirement for HIV-2 genome dimerisation.

On the contrary, mutation of the palindrome located in the loop of SL-1 (putative DIS) had no effect on genomic RNA dimerisation, in spite of abrogating *in vitro* dimerisation of a short RNA transcript and presenting similarities with the HIV-1 DIS. Previous *in vitro* studies on HIV-2 dimerisation had proposed that the putative DIS mediated dimerisation (Dirac et al., 2001; Lanchy and Lodmell, 2002). Here, we have shown that discrepancies exist between *in vitro* and *in virio* data and that, in the context of the full-length genome, the putative DIS palindrome was dispensable for HIV-2 RNA dimerisation. A similar disparity has also been reported in HIV-1 (Sakuragi and Panganiban, 1997). The differences observed between the *in vitro* and *in virio* data may be due to the presence of *trans*-acting factors such as the NC protein, which has been shown to enhance dimerisation of HIV-1 and HIV-2 RNA transcripts *in vitro* (Darlix et

al., 1990; Dirac et al., 2002; Lanchy et al., 2003b) and induce a change of conformation, called 'maturation', of HIV-1 RNA dimers *in vitro* (Feng et al., 1996b; Mujeeb et al., 2007; Muriaux et al., 1996a; Takahashi et al., 2001). In addition, the viral protease PR has been reported to be essential for the maturation of the dimer following particle budding in both HIV-1 and MLV and the stability of dimer isolated from PR^- viruses was found to be similar to that observed in immature particles (Fu et al., 1994; Fu and Rein, 1993; Shehu-Xhilaga et al., 2001b; Song et al., 2007b). Furthermore, several long-range interactions have been described in the HIV-1 and HIV-2 leader RNAs (Abbink and Berkhout, 2003; Dirac et al., 2002; Huthoff and Berkhout, 2001; Paillart et al., 2002) and some of them have been implicated in the regulation of dimerisation (Abbink et al., 2005; Dirac et al., 2002; Lanchy et al., 2003b).

Interestingly, HIV-1 with a mutated DIS was found to contain dimeric genomes, yet this virus had a delayed replication kinetic in SUP-T1 cells and a reduced level of dimeric genomes packaged (Berkhout and van Wamel, 1996). Several studies confirmed that the DIS was required for HIV-1 infectivity in T-cells (Laughrea et al., 1997; Paillart et al., 1996a), however, this requirement appeared to be cell-type dependent, as PBMCs could support replication of DIS mutants over the course of 40 days (Hill et al., 2003). A similar cell-type dependency might exist in HIV-2 and it would be interesting to assess whether it is the case. Furthermore, one might envisage that the palindrome *pal*, as the main dimerisation site in HIV-2, would act like the HIV-1 dimerisation site and therefore be required for replication in a cell-type dependent manner. This hypothesis has not yet been explored. However, the replication of the dimerisation-deficient viruses in T-cells was assessed and it was found that these viruses could not replicate. Several hypotheses could explain such a replicative defect. The mutations introduced might affect the translation of the Gag polyprotein, possibly by introducing an additional

structure in the leader RNA and further reducing the efficiency of ribosomal scanning, as has been shown in HIV-1 (Miele et al., 1996). However, it was suggested that HIV-2 could translate the Gag polyprotein using an IRES located downstream of the Gag authentic start codon at position 545 of HIV-2_{ROD} (Herbreteau et al., 2005). This would therefore not be affected by the mutations introduced in the RNA leader. Gag polyprotein synthesis and its proteolytic cleavage were analysed at 48 h post-transfection by western blot analysis and were found to be unaffected by the mutations. This does not exclude the possibility that the mutations could affect Gag synthesis at earlier time points. Moreover, it has not been fully evaluated whether the sequence and ratio of the proteolytic processing of the Gag polyprotein might have been disturbed. The proteolytic processing of the Gag protein as well as the ratio of Gag/Gag-Pol have been shown to influence RNA dimerisation in HIV-1 (Shehu-Xhilaga et al., 2001a; Shehu-Xhilaga et al., 2001b). The reverse is also true and mutations near the dimerisation signal of HIV-1 were found to result in the delayed processing of the p2 peptide from the CA-p2 precursor (Liang et al., 1999a).

All non-replicative mutant viruses showed a dimerisation and encapsidation defect. The packaging defect alone could be responsible for the reduction in infectivity as it would result in the release of more empty particles. The reduction in the proportion of dimeric genomes in the mutant virions could also impair viral replication, as seen in HIV-1 and MLV, by affecting reverse transcription and reducing the recombination rate, for example by preventing the use of the alternative template in intermolecular first-strand transfer or, in the case of strand breakage, in intramolecular first-strand transfer (Berkhout et al., 1998; Chin et al., 2008; Hu and Temin, 1990a; Jones et al., 2008; Mikkelsen et al., 2000; Mikkelsen et al., 2004; Moore et al., 2007; Paillart et al., 1996a; Panganiban and Fiore, 1988; Shen et al., 2000). The efficiency of reverse transcription

in the dimerisation-deficient viruses was assessed by measuring the level of strong-stop cDNA synthesis. However, due to technical difficulties, the assay was inconclusive and it was not possible to determine whether this stage of reverse transcription was affected. It would be interesting to further investigate this aspect of the virus life cycle and determined whether the initial stages of reverse transcription are affected when the RNA genomes are less dimeric, and in particular to measure the efficiency of the first-strand transfer.

Since the *cis*-acting signals involved in dimerisation and encapsidation are in close proximity, it was difficult to discriminate if one or both defects are responsible for inhibiting viral replication. To determine if reduced genome encapsidation restricted the replication of the Psi mutant, its packaging defect was accounted for by increasing the amount of genome targeted onto the cells in an infectivity assay. Interestingly, this did not restore the infectivity of the mutant. Similarly, reducing the amount of wild type genomes to the same level as the mutant virus had no impact on wild type HIV-2 replication. Altogether, these data implied that the reduction in viral RNA encapsidation was not the sole factor responsible for the fall in viral infectivity observed with the Psi mutant. Furthermore, it seemed unlikely that a two- to threefold reduction in genome encapsidation could completely abrogate viral replication. Similarly, a two- to fourfold decrease in genome dimerisation alone might not be sufficient to inhibit viral replication. Hence, it appeared probable that a combination of defects, including - but not exclusively - dimer formation and genome packaging, were responsible for the block in viral replication.

It has also been suggested that the presence of the RNA genome may contribute to particle morphogenesis (Campbell and Rein, 1999; Campbell and Vogt, 1995; Muriaux et al., 2001). This would imply that the Psi mutants, which packaged less RNA, would

produced more empty particles or particles with an aberrant morphology than the wild type virus. This could trigger the replication defect observed. This hypothesis has been tested and will be detailed in Chapter 5.

Nonetheless, from these data emerged a correlation between RNA dimerisation, packaging and viral replication. Such a link has been observed in HIV-1 (reviewed in Greatorex, 2004; Paillart et al., 2004b) and is supported by the close proximity of the elements required for RNA dimerisation and packaging. Furthermore, a recent study of the HIV-2 packaging signal proposed that genome encapsidation and viral replication required the presence of an extended SL-1, as depicted in Figure 3.1B, that includes the formation of a lower stem, termed stem B (Lanchy and Lodmell, 2007). The DM and Pal mutations of the packaging signal both disrupted the SL-1 structure, including stem B, and this alteration of the SL-1 structure might explain the inhibition of viral replication observed. However, in the SM2 and SM3 mutants, which also failed to infect and replicate in T-cells, stem B remained unaffected by the mutation and it appeared that SL-1 was further extended due to the substitution of the GGAG residues at position 392-395 (see Figure 3.3). The structural context of the dimerisation and packaging elements and its implication for HIV-2 infectivity has been further explored and is detailed in Chapter 4.

4 Sequence and structural context of stem-loop 1 and its implication for HIV-2 replication

4.1 Introduction

It has been demonstrated that the 5′ RNA leader of HIV-1, comprising the RNA elements required for genome dimerisation and encapsidation, is highly structured (Baudin et al., 1993; Clever et al., 1995; Harrison and Lever, 1992; McBride and Panganiban, 1996; Sakaguchi et al., 1993). As part of the encapsidation process, the NC domain of the Gag polyprotein recognises and binds to specific motifs within the HIV-1 RNA leader via its two zinc fingers (Aldovini and Young, 1990; Gorelick et al., 1990), and the exact nature and affinity of this binding has been extensively studied using a range of methods (Amarasinghe et al., 2000a; Amarasinghe et al., 2001; Bacharach and Goff, 1998; Berkowitz and Goff, 1994; Berkowitz et al., 1993; Clever et al., 1995; Damgaard et al., 1998; Dannull et al., 1994; De Guzman et al., 1998; Hagan and Fabris, 2003; Zeffman et al., 2000). Specific RNA structures are also critical for genome dimerisation. The position of the HIV-1 DIS in a stem-loop structure favours loop-loop interaction and formation of the 'kissing-loop' complex that was shown to be essential for dimerisation (Clever et al., 1996; Haddrick et al., 1996; Laughrea and Jette, 1994, 1996; Mujeeb et al., 1998; Muriaux et al., 1996b; Paillart et al., 1994; Paillart et al., 1996b; Paillart et al., 1997; Skripkin et al., 1994). As for HIV-1 packaging, NC has been proposed to enhance HIV-1 RNA dimerisation by promoting maturation of the RNA dimer from a weak dimer to a more stable form (Feng et al., 1996b; Mujeeb et al.,

2007; Muriaux et al., 1996a; Takahashi et al., 2001). Protease-dependent proteolytic processing of the Gag protein appears to be required for this 'maturation' of the dimer (Fu et al., 1994; Shehu-Xhilaga et al., 2001b; Song et al., 2007b), although it has been suggested that Gag processing alone is not sufficient (Fu et al., 2006).

In HIV-1, it has been shown that the structural context of stem-loop 3 (SL-3) and SL-1 is important for HIV-1 genomic RNA encapsidation (Clever and Parslow, 1997). Similarly, altering the structural context of the HIV-1 DIS, in particular a lower stem structure part of SL-1 and termed stem B, has a critical effect on dimerisation (Clever and Parslow, 1997; Clever et al., 1996; Laughrea et al., 1997; Laughrea et al., 1999; Muriaux et al., 1996b; Russell et al., 2003a; Sakuragi et al., 2003).

The HIV-2 RNA leader has been shown to contain a stem-loop structurally equivalent to HIV-1 SL-1 (Baig et al., 2007; Damgaard et al., 1998; Dirac et al., 2001; Lanchy et al., 2004). Interestingly, long-term culture and phenotypic reversions of *pal*-mutated HIV-2 have suggested the existence of an extended SL-1, including formation of a lower stem structure (stem B), and it was suggested that the integrity of this stem is necessary for HIV-2 replication and RNA packaging (Lanchy and Lodmell, 2007).

The 3′ end of *pal* is involved in the formation of stem B. However, mutation of the 5′ end of *pal* (SM2), which is not part of stem B, significantly reduced infectivity and impaired viral replication (see Chapter 3). Hence, the precise sequence and structural elements required for HIV-2 replication are not yet resolved.

In this study, we aimed to clarify this and to identify the sequence and structure requirements for the individual processes of HIV-2 genome dimerisation, packaging and viral infectivity. Using a mutagenesis approach, the SL-1 of the HIV-2 leader RNA was mutated to disrupt the previously implicated motifs, *pal* and stem B. Compensatory mutations were also created to distinguish the relative importance of the structural

elements over nucleotide sequence alone. A GGAG motif corresponding to the 5′ end of the palindrome *pal* was shown to be critical for both genome dimerisation and viral replication. Surprisingly, RNA encapsidation could be affected independently of genome dimerisation. However, it was found that dimerisation-deficient viruses fail to replicate in T-cells. The data obtained suggested that stem B was not required for viral infectivity and clearly identified the GGAG motif at position 392-395 as essential for HIV-2 replication.

4.2 Role of the SL-1 sequence and structural motifs in genomic RNA dimerisation and encapsidation

4.2.1 Stem-loop 1 mutants

To investigate the importance of the postulated structure termed stem B, located at the base of SL-1 and involving part of the Psi palindrome *pal*, a series of substitution mutants (SM4 to SM9; Figure 4.1) that affected the formation of the stem B structure and the bulge of SL-1 were created (Figure 4.2A). The SM4 and SM5 mutations both led to the disruption of stem B (Figure 4.2B and C), with SM4 also disrupting *pal* while SM5 retained a palindromic sequence (Figure 4.1A). In SM6, the 3′ end of SL-1 was mutated to disrupt stem B but leaving *pal* unaffected (Figure 4.1A and B, residues 439-442 shown in green; Figure 4.2D). To assess the relative importance of the stem B structure compared to its constituent sequences, a double mutant of SM4 and SM6 (SM9) was created (Figure 4.1A). This mutant restored a stem B structure using alternative sequences (Figure 4.2G).

There are several GGAG motifs present in the palindrome *pal* and stem B. The previously described SM2 was a substitution of the GGAG motif at position 392-395 of

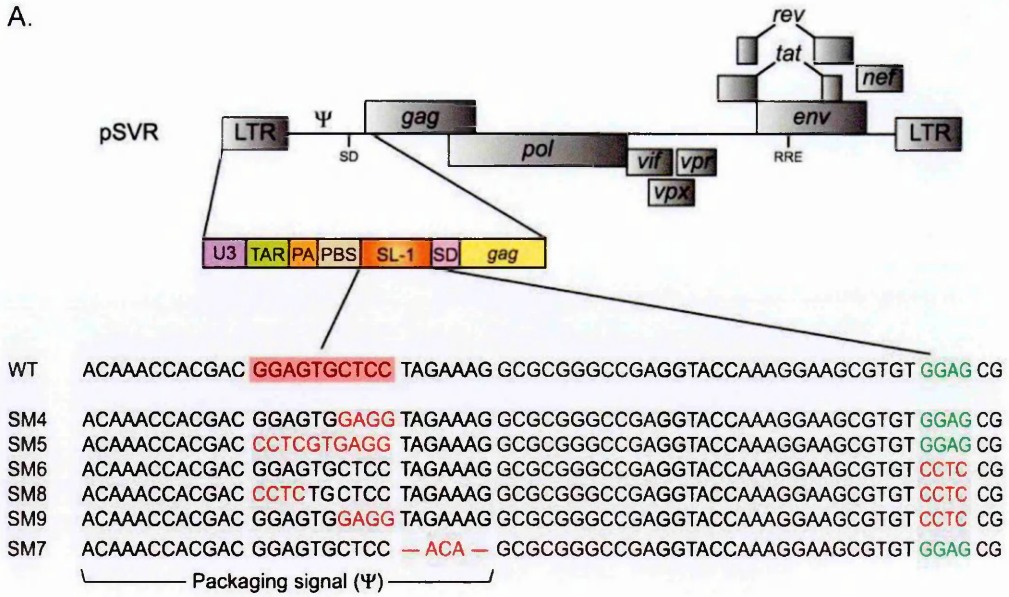
the HIV-2_{ROD} genome and caused a dramatic decrease in viral infectivity and genome dimerisation (see Chapter 3). SM6 affected the second GGAG motif, at the 3′ end of SL-1 (Figure 4.1B). Similar purine tetrads have been implicated in RNA packaging and Gag binding in both HIV-1 (Aldovini and Young, 1990; Clavel and Orenstein, 1990; Damgaard et al., 1998; Lever et al., 1989; Shubsda et al., 2002) and HIV-2 (Damgaard et al., 1998; Garzino-Demo et al., 1995; McCann and Lever, 1997; Poeschla et al., 1998), hence a double mutant of SM2 and SM6 was created to analyse the effect of substitutions in both GGAG motifs (SM8; Figure 4.1A and Figure 4.2F).

Computer generated predictions and biochemical studies (Baig et al., 2007; Damgaard et al., 1998; Dirac et al., 2001; Lanchy et al., 2004) all suggested that SL-1 was a discontinuous helix (see Figure 4.1B). Thus, the nucleotides involved in the bulge of SL-1 (position 402-409) were substituted in SM7 (Figure 4.1A), resulting in the formation of an extended and highly stable SL-1 (Figure 4.2E).

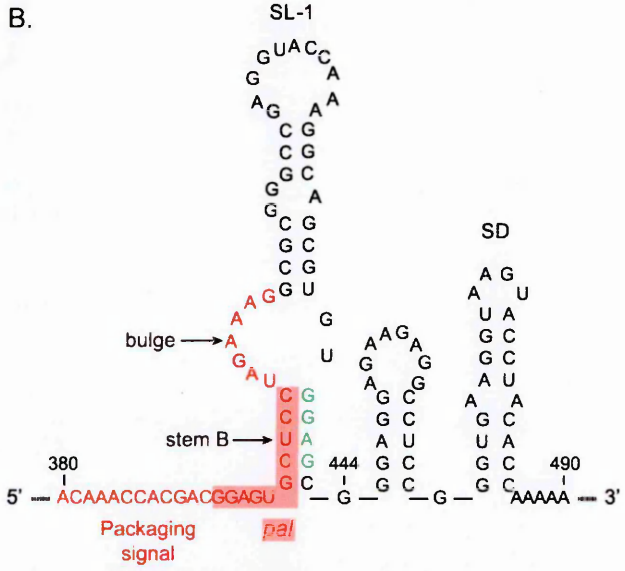
The effects of each mutant on stem B, the palindrome *pal* and the bulge structure are summarised in Figure 4.1C.

To ensure that the mutation introduced did not severely affect protein production, the level of viral proteins in transfected Cos-1 cells and newly released virions was assessed by western blotting using a panel of HIV-2 positive patient sera (Figure 4.3). At 48 h post-transfection, the level of the Env glycoproteins gp125 and gp36 appeared maintained across the range of mutants. The level of RT (p66), Gag (p57) and proteolytic cleavage products MA-CA-p2 (p41) and CA (p26) was equivalent in the wild type and mutant transfected cells and subsequently released particles. The slight reduction in the level of p57 Gag observed in the SM7 virions was not consistent in between experiments.

A.



B.



C.

	Pal	Stem B	Bulge
WT	+	+	+
SM4	-	-	+
SM5	+	-	+
SM6	+	-	+
SM8	-	-	+
SM9	-	+	+
SM7	+	+	-

Figure 4.1: Genomic and structural context of the stem-loop 1 mutations.

(A) Genomic organisation of the HIV-2_{ROD} molecular clone (pSVR) and localisation of the SL-1 mutations. The main functional elements of the 5' leader region are indicated: TAR, Tat response element; PA, PolyA signal; PBS, Primer binding site; SL-1, Stem-loop 1; SD, Splice donor. The main packaging determinant Psi (Ψ) is indicated. WT, wild type; SM4 to SM9, Substitution Mutant 4 to 9. The palindrome *pal* (392-401) is denoted by the red shadowing in the WT sequence and the 3′ end of SL-1 involved in the formation of stem B (439-442) is shown in green. The mutated nucleotides are indicated in red. All numberings refer to the start of the HIV-2_{ROD} genome. (B) Structure of WT SL-1 as determined by biochemical analysis (Baig et al., 2007; Damgaard et al., 1998; Dirac et al., 2001; Lanchy et al., 2004) and free-energy minimisation using *mfold*. The packaging signal Psi (380-408) and *pal* (392-401) are shown in red and shadowed red respectively. The residues involved in the formation of stem B (439-442) appear in green. The bulge mutated in SM7 is indicated. (C) Presence (+) or absence (-) of *pal* (392-401), stem B (397-401/439-442) and bulge (402-408) in the WT and mutated viruses.

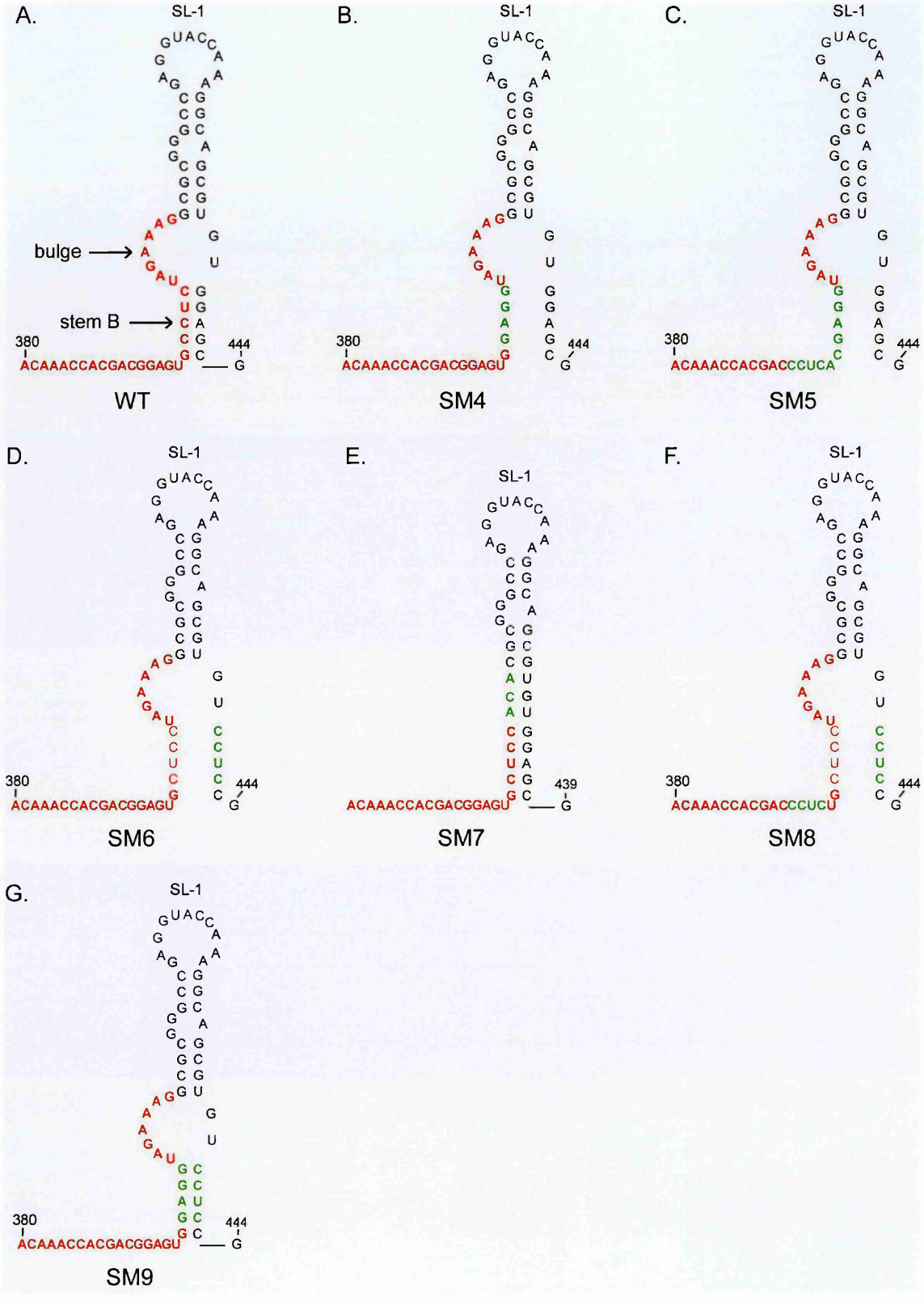


Figure 4.2: Predicted structures of the SL-1 mutants.

Predicted structure of SL-1 (nt 380-444 of HIV-2_{ROD}) of wild type (WT; A), SM4 (B), SM5 (C), SM6 (D), SM7 (E), SM8 (F) and SM9 (G) using the algorithm *mfold*. The Psi region is indicated in red and the mutated residues in green.

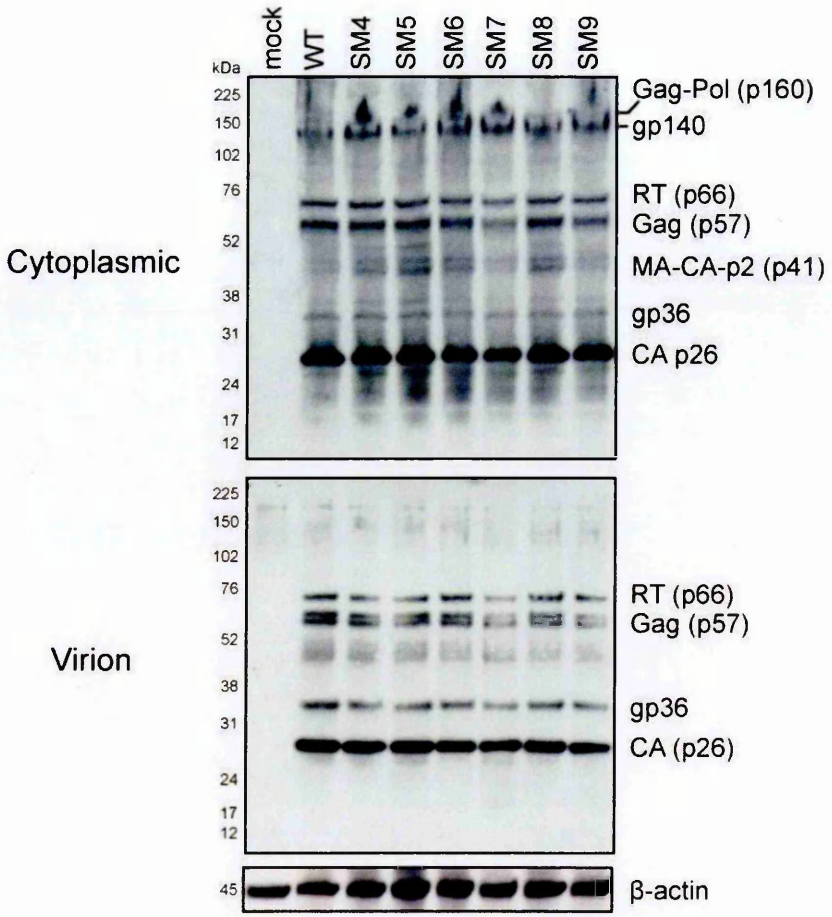


Figure 4.3: Gag protein level in wild type and mutant HIV-2.

Cos-1 cells were transfected with wild type (WT) and SM4 to SM9 mutant proviral constructs. Cells (top panel) and virions (middle panel) were harvested 48 h later and analysed by western blot using HIV-2 positive patient sera (ARP501, NIBSC; 1:1000) and a rabbit anti-human-HRP antibody (Dako; 1:1000). An equal volume of the cytoplasmic fraction (15 µl) was used and the virion fraction was normalised on the RT activity of the virus preparation. Equal loading of the cytoplasmic protein fraction was verified using a mouse anti-actin antibody (Abcam; 1:5000, bottom panel) and a goat anti-mouse-HRP antibody (Dako; 1:1000). WT, wild type; mock, mock transfected cells. The sizes of the protein markers (full-range Rainbow; GE Healthcare) are indicated on the left. The position of the Gag-Pol polyprotein (p160), Env gp140 and gp36 proteins, RT (p66), Gag polyprotein (p57) and Gag cleavage products MA-CA-p2 (p41) and CA (p26) is indicated on the right. The multiple bands seen for the Gag protein represent the different isoforms (p57, p50 and p44) depending on the initiation codon usage, as described by Herbreteau et al. (2005).

4.2.2 Substitution of the GGAG motif at position 392-395 negatively affects genome dimerisation

Previous data link elements located within the HIV-2 packaging signal with *in vitro* dimerisation (Lanchy et al., 2003b). In addition, the results obtained in the present study showed that the HIV-2 Psi region was involved in genomic RNA dimerisation *in virio* and that it was associated with viral infectivity. Hence, the ability of the SL-1 mutant viruses to dimerise *in virio* was analysed (Figure 4.4). Wild type and mutant virions were purified following transfection of Cos-1 cells and the genomic viral RNA was analysed by native northern blot, with RNA inputs normalised on the RT activity of the virion preparation.

As shown in Figure 4.4, disruption of stem B by mutation of the 3′ end of *pal* in SM4 or by mutation of the GGAG motif at position 439-442 in SM6 resulted in decreased genome dimerisation (left panel). However, on a longer exposure of the blot (Figure 4.4, right panel), dimeric genomes are still easily identifiable, indicating that genome dimerisation could occur in the absence of stem B. Interestingly, SM9, a combination of SM4 and SM6 which restored stem B but disrupted *pal*, had a high level of dimeric genomes (Figure 4.4), suggesting that *pal* was dispensable for dimerisation; yet, mutations of *pal* resulted in a decrease in dimerisation in the Pal and SM2 mutants (see Figure 3.4). When comparing the dimerisation of the two *pal* substitution mutants, SM2 and SM4 (Figure 3.4 and Figure 4.4), with that of the *pal* deletion mutant Pal (Figure 3.4), it appeared that the 5′ end of *pal* (residues 392-395) was essential. Indeed, mutation of the 5′ end of *pal* in SM2 triggered the same dimerisation defect as deletion of the whole palindrome (Pal), whereas mutation of the 3′ end of *pal* in SM4 (residues 397-401) had a less marked reduction in the level of dimeric genome present in the virions.

Analysis of SM5 confirmed that a palindromic sequence in Psi was not a critical factor for dimerisation as this mutant contained a substitute palindromic sequence at the same position as wild type *pal* but failed to dimerise (Figure 4.4).

SM8, a combination of SM2 (nt 392-395) and SM6 (nt 439-442), also failed to dimerise (Figure 4.4). Since SM2 did not dimerise (Figure 3.4), but SM6 could, albeit at a lower efficiency than wild type, the GGAG motif at position 392-395 appears to be dominant for genome dimerisation.

SM7 dimerised almost to wild type level, indicating that the bulge in SL1 is dispensable for genome dimerisation (Figure 4.4). The minor defect observed may be due to a more stable stem impairing the transition from the initial duplex to the extended dimer.

These data, together with the previous analysis of the SM2 mutant, indicated that neither stem B, nor merely a non specific palindromic sequence in the Psi region, was critical for genome dimerisation. The sole factor in common between the mutants which failed to dimerise was the absence of the GGAG motif at position 392-395 (see Figure 4.1C and Figure 3.3).

4.2.3 The reduction in genome dimerisation does not correlate with a packaging defect

Genome dimerisation and packaging are frequently described as interdependent (for review see Russell et al., 2004). In HIV-2, the main packaging determinant (Griffin et al., 2001; McCann and Lever, 1997) overlaps with sequences shown to be involved in RNA dimerisation *in vitro* (Baig et al., 2007; Lanchy et al., 2003b) and *in virio* (present study). In order to further investigate this relationship in HIV-2, the packaging efficiencies of the SL-1 mutants were measured by RPA (Figure 4.5).

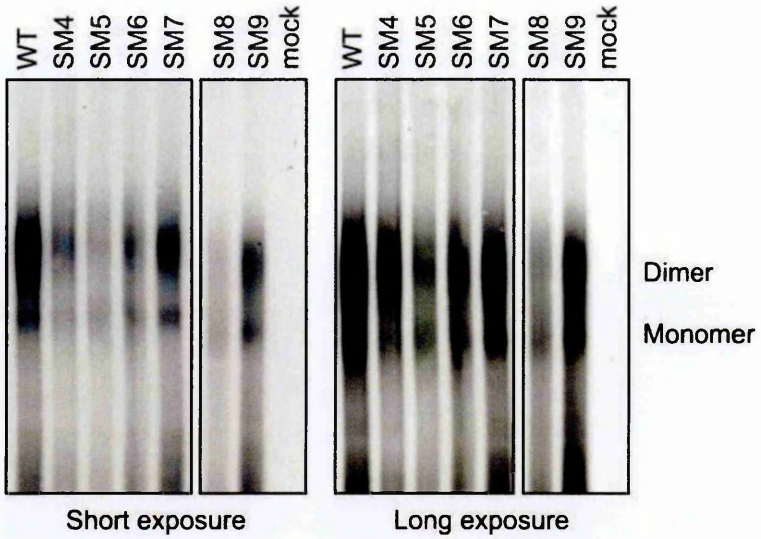


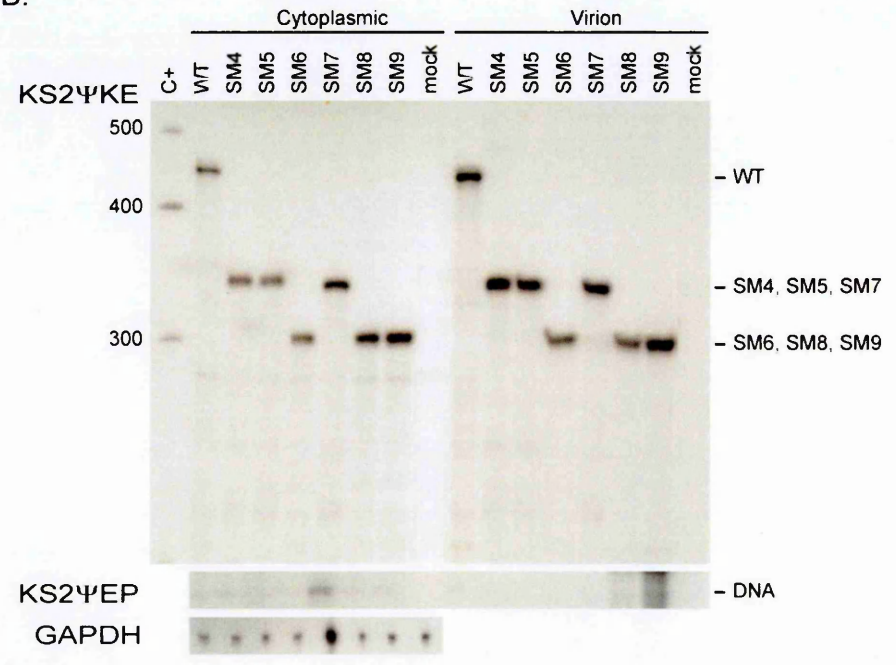
Figure 4.4: Dimerisation of the HIV-2 genome requires the presence of a GGAG motif at position 392-395.

Native northern blot analysis of wild type (WT) and mutant genomic RNA. WT and mutant proviral constructs were transfected into Cos-1 cells and virions were harvested after 48 h. The genomic RNA was normalised by RT activity and analysed by native northern blot. The right panel shows a longer exposure (15 minutes) of the blots presented in the left panel (1 minute exposure). mock: mock transfected cells. Representative blot from three independent experiments is shown.

A.

	Size of genomic RNA protected fragments using KS2ΨKE (nt)
WT	445
SM4	349 + 91
SM5	349 + 85
SM6	308 + 132
SM7	341 + 95
SM8	308 + 86
SM9	308 + 91

B.



C.

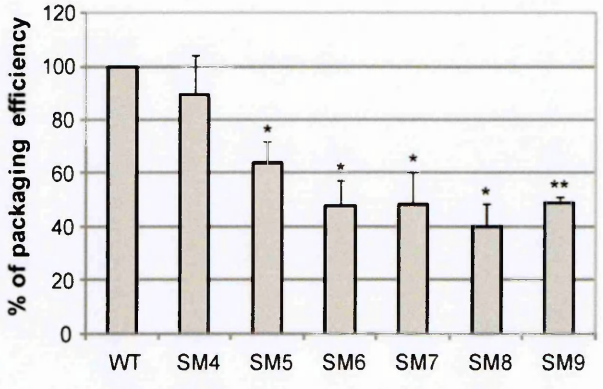


Figure 4.5: Packaging efficiencies of the SL-1 mutants.

(A) Size of the genomic RNA fragments protected by the KS2ΨKE riboprobe. (B) Representative RPA of wild type (WT) and mutant HIV-2 RNAs. Proviral constructs were transfected into Cos-1 cells and virions were harvested after 48 h. Cytoplasmic and virion RNAs were extracted and subjected to an RPA with riboprobes KS2ΨKE, KS2ΨEP and GAPDH. 1 µg of cytoplasmic RNA and an equivalent of 2500 cpm of virion RNA, as measured by RT assay, were used. C+, Century Plus RNA markers (Ambion). (C) Packaging efficiencies of the mutant viruses relative to WT. The average and sd of three independent experiments is shown. Significant difference relative to WT packaging efficiency was assessed using a two-tailed unpaired Student t test. *, p value < 0.05; **, p value < 0.001.

The protected fragments generated by hybridisation of the KS2ΨKE riboprobe to wild type and mutant genomic RNAs are indicated in Figure 4.5A and a representative RPA using the KS2ΨKE riboprobe is shown in Figure 4.5B. The level of cytoplasmic RNA was adjusted by probing for GAPDH and plasmid DNA contamination was accounted for using the KS2ΨEP riboprobe, which recognises sequences between the U3 region and the PBS. Packaging efficiencies relative to wild type are presented in Figure 4.5C.

Contrary to what was observed with the Psi mutants, there was no interdependent relationship between the ability of the genome to dimerise and the efficiency with which it was packaged. While the dimerisation-competent SM4 mutant packaged as efficiently as the wild type virus, the SM6, SM7 and SM9 mutants packaged at approximately 50 % of wild type, despite being able to dimerise. Similarly, the dimerisation-deficient mutants SM5 and SM8 both showed reduced packaging, even though the intensity of the defect was variable. These data suggest that the HIV-2 genome dimerisation and encapsidation processes may not be as interdependent as initially thought. It is interesting to note that there seems to be a nonreciprocal relationship between genome dimerisation and packaging in HIV-2. While all mutants that failed to dimerise showed a packaging defect, the reverse is not true. Therefore, one might conclude that RNA dimerisation is important for encapsidation and that the GGAG motif at position 392-395 is required; however, there must be an alternative mechanism to explain the case of those mutants with a normal level of RNA dimer but a reduced RNA content.

Although these results did not corroborate the 'apparent' level of genomic RNA present in the virions analysed by northern blot (Figure 4.4), the RPA remained the most appropriate and sensitive method to calculate packaging efficiencies. The sensitivity of the technique is reported to be at least ten times greater than for northern blot (Ambion technical bulletin #500) and most importantly, the level of viral RNA present in

the cytoplasm of the cells as well as the contamination of the sample by plasmid DNA are taken into account. In addition, breakage of the genomic RNA has little impact on an RPA as it is based on the detection of short protected fragment as compared to full-length genome of approximately 9500 nt in the northern blot. Finally, an RPA allows the visualisation and quantification of any spliced RNA being incorporated into virions. The transfection of Cos-1 cells results in the over-expression of the viral genomes and one might envisage that this high-level expression might suppress the effect of otherwise disruptive mutations, as suggested previously (Luban and Goff, 1994).

4.3 Effect of the mutations on viral replication and infectivity

4.3.1 Mutations of the 392-395 GGAG motif reduce infectivity and impair viral replication in T-cells

Mutations within the HIV-2 packaging signal and SL-1 region have been reported to affect HIV-2 replication (Griffin et al., 2001; Lanchy and Lodmell, 2007; McCann and Lever, 1997). To determine whether the dimerisation-deficient SL-1 mutants were infectious, a single-round infectivity assay using the GHOST reporter cell line was performed. Increasing amounts of virus, normalised on RT activity, were inoculated onto the cells and the percentage of infected cells was determined by FACS analysis. The results obtained for the highest virus input were representative of the data set and are shown in Figure 4.6A. Disruption of stem B in SM4 and SM6 resulted in a significant decrease in the percentage of infected cells, while the SM9 double mutant appeared fully infectious, suggesting that stem B may be important for infectivity. Analysis of the SM5 and SM8 mutants showed a correlation between the ability to dimerise and viral infectivity. These SM5 and SM8 mutants failed to dimerise and

demonstrated the most striking reduction of infectivity, at least fivefold lower than wild type, while SM4 and SM6, which had a small decrease in virion dimers, showed a less dramatic reduction of viral infectivity (Figure 4.6B). It is interesting to note that SM8, which lacked two GGAG motifs, showed the strongest decrease in infectivity, suggesting that these purine-rich motifs may be important for viral replication. On the other hand, dimerisation-competent mutants SM7 and SM9 were fully infectious (Figure 4.6A and B).

To confirm these findings in a system allowing viral replication, PM1 T-cells were infected with wild type and mutant viruses and the replication kinetics of the viruses were followed by measuring the RT activity in the culture supernatants (Figure 4.6C). The correlation between genome dimerisation and viral replication was clearly reproduced. Replication patterns for the SM4, SM7 and SM9 viruses were similar to that of wild type virus, demonstrating that both stem B and the SL-1 bulge are dispensable for HIV-2 replication. The SM6 virus was replication-competent but showed a delay, which could be explained by the reduced infectivity observed in the single-round infectivity assay and correlated with a small decrease in RNA dimerisation and packaging. This result suggested that the second GGAG motif may be involved in viral replication, possibly acting as a Gag binding site. Finally, the SM5 and SM8 viruses did not replicate detectably in PM1 cells over a 15 to 20 day monitoring period, suggesting that the original palindromic sequence *pal* but not the alternative palindrome may be essential for viral replication. Comparison of the replication fitness of SM5 and SM8 with the previously analysed SM2 and Pal mutants highlighted the 5′ end of *pal* as being required for viral replication.

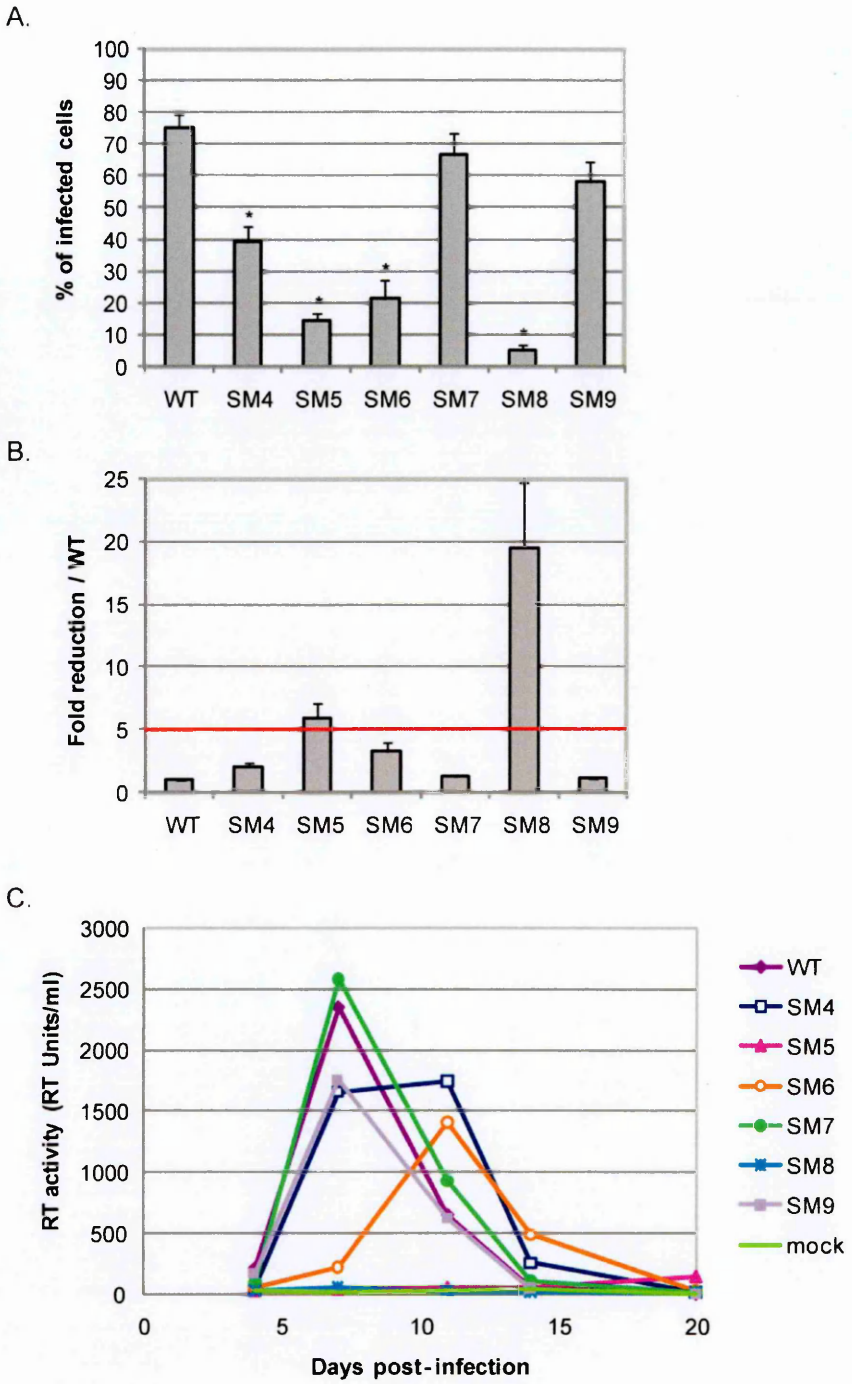


Figure 4.6: The infectivity of the viruses correlates with their ability to dimerise their genome.

(A) Single-round infectivity assay. 2000 cpm of virus, as measured by RT activity, was used to infect 2×10^5 cells. Infected cells were harvested 72 h later and GFP expression was analysed by FACS. WT, wild type. The average and sd of four independent experiments is plotted.*, two-tailed unpaired Student t test p value < 0.01. (B) Viral infectivity from (A) expressed as fold reduction in infectivity for the mutant viruses compared to WT virus. A fivefold reduction over WT is indicated by the red line. Average and standard error of at least four experiments are shown. (C) Replication kinetics of WT and mutant viruses in PM1 T-cells. 10000 cpm of virus, as measured by RT activity, was used to infect 1×10^6 cells and replication followed by measuring RT activity in the culture supernatant every 4 days. Mock, mock infected cells.

In particular, and again consistent with the capacity to dimerise the genome, the presence of the GGAG motif at position 392-395 appeared to be critical for viral replication, being present in all replication-competent viruses and absent in all non-infectious viruses mutants (Table 3). This was further supported by phylogenetic analysis, which revealed that this motif was 100 % conserved in all HIV-2 subtypes and most of the related SIVs, namely SIV_{SM}, SIV stump-tailed macaque (SIV_{STM}), SIV pig-tailed macaque (SIV_{MNE}) and SIV_{MAC}, as shown on Figure 4.7 (HIV sequence compendium 2008).

Table 3: Correlation between viral infectivity and the sequences and structural elements of the HIV-2 leader.

Virus	**Infectivity**[a]	Sequence elements		Structural elements				
		Psi Palindrome[b]	**GGAG (392-395)**	Stem B	Psi 1[c]	Splice donor[c]	Psi 2[c]	Psi 3[c]
WT	+	+	+	+	+	+	+	+
SM4	+	-	+	-	+	+	+	+
SM6	+	+	+	-	-	+	+	+
SM7	+	+	+	+	+	+	+	+
SM9	+	-	+	+	+	-	+	+
SM2	-	-	-	+	-	+	-	+
SM5	-	+	-	-	+	+	-	-
SM8	-	-	-	-	-	-	+	+
Pal	-	-	-	-	+	+	+	+

[a], Replication kinetics in PM1 T-cells (measure of RT activity) based on 3 independent experiments.

[b], Palindrome *pal* at position 392-401 of the $HIV\text{-}2_{ROD}$ genome.

[c], Stem-loop structures previously described by Dirac and co-workers (Dirac et al., 2001) and predicted by the *mfold* program.

HIV-2 isolate	Accession number		Sequence	
ROD	M15390	380	ACAAACCACGACGGAGTGCTCCTAGAAAG	408
MCR35	AY509260	381	.....	409
MCN13	AY509259	381	.....	409
BEN	M30502	936	.....	964
CRIK147	DQ307022	682	.....	710
ALI	AF082339	931	.....	959
SBLISY	J04498	381	.....	409
PEI2	U22047	904	.....	932
NNVA	EU980602	930	.....	957
MDS	Z48731	380	.....	407
D194	X52223	381	.....	408
UC2	U38293	935	.....	962
STGH1	M31113	382	.....	409
96FR12034	M30895	376	.....	403
Abt96	AY530889	427	.....	449
CAM2CG	AF208027	307	..T.....	332
	D00835	1136	.....	1157

Figure 4.7: Phylogenetic analysis of the HIV-2_{ROD} packaging signal.
The packaging signal of HIV-2_{ROD} (nt 380-408) was aligned against sequences of fifteen other HIV-2 isolates using the basic local alignment search tool or BLAST (Altschul et al., 1997; http://blast.ncbi.nlm.nih.gov/Blast.cgi). The GGAG motif at position 392-395 is indicated by the red box. The name and GenBank accession number of each isolate is indicated.

4.3.2 Partial characterisation of revertant viruses

Mutations of SL-1 that affected the GGAG motif at position 392-395 of the HIV-2_{ROD} genome resulted in an absence of detectable viral replication in PM1 T-cells as measured by the RT activity of the culture supernatant (see Figure 4.6). Some of the mutations introduced only affected a few nucleotides, leaving the possibility for the virus to revert. Yet, no reversion event was observed in the first 20 days of the experiment. In order to analyse whether the replication of some mutant viruses could be rescued by a reversion of the mutation or a second site mutation, the replication of SM5 and SM8 in PM1 cells was followed for a longer period of time (Figure 4.8A). The replication of the DM and Pal mutants was also measured, as PM1 cells were shown to support HIV-2 replication more efficiently than Jurkat cells (see Figure 3.6) and the absence of reversion event in Jurkat cells might have been due in part to a low replication efficiency of the virus.

Interestingly, while RT activity could not be detected for DM after 30 days, the SM5, SM8 and Pal viruses showed detectable viral replication when cultured for more than 20 days in PM1 cells (Figure 4.8A). In order to determine whether the replication in a particular culture was due to a reversion of the mutation or was simply a delayed replication phenotype, cell-free culture supernatant was added to fresh PM1 cells and the replication of the viruses was followed by measuring the RT activity in the culture supernatant (Figure 4.8B). The passage of each virus was performed at day 25 post-infection, when the RT activity detected was significantly above background level. Unfortunately, this time point did not correspond to the peak of infection for Pal and SM5.

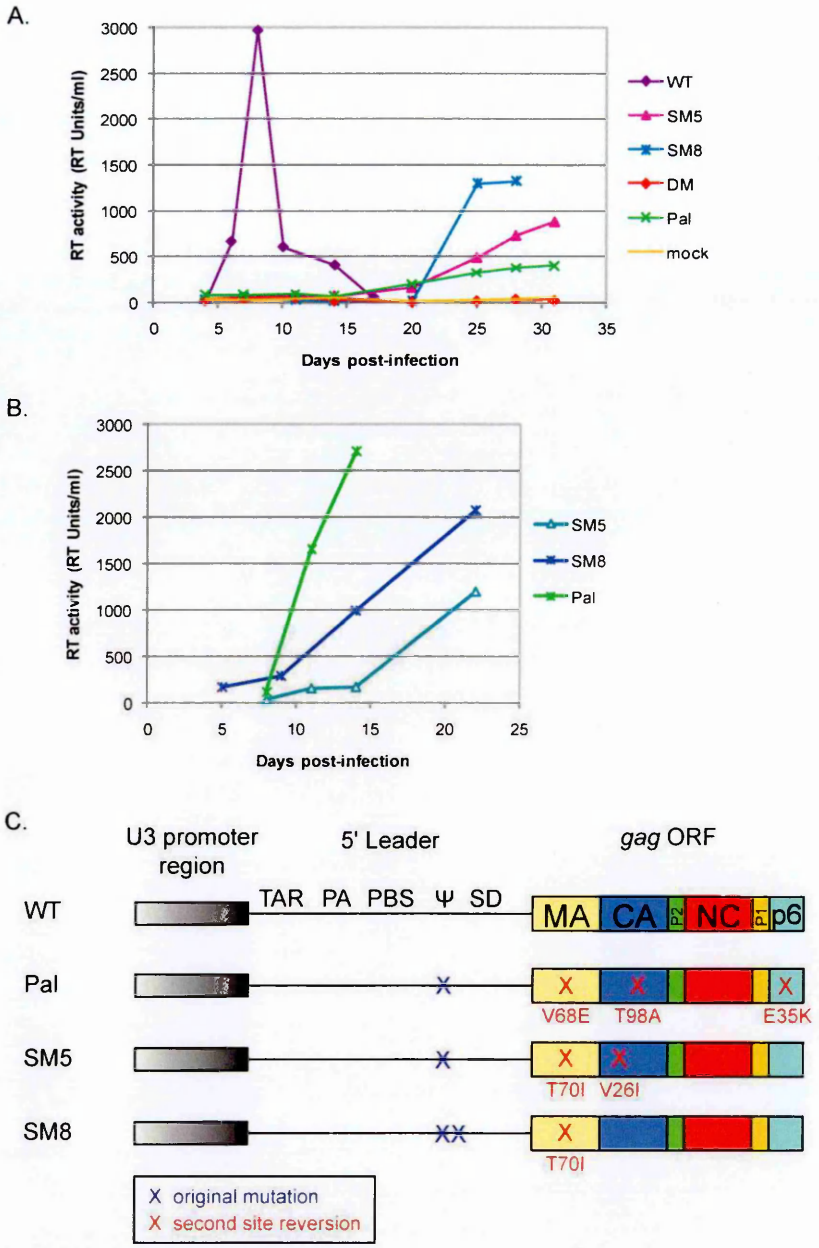


Figure 4.8: Analysis of the revertant viruses reveals second site mutations in Gag.
(A) 10000 cpm of wild type (WT) and mutant viruses, as measured by RT activity, was used to infect 1 × 10^6 PM1 cells and replication followed by measuring the RT activity in the culture supernatant every 2 to 4 days. Mock, mock infected cells. (B) At day 25 post-infection, Pal, SM5 and SM8 viruses were passaged onto fresh cells. 1 ml of culture was spun at 300 × *g* for 5 min and 500 µl of cell-free supernatant was added to 1 × 10^6 PM1 cells. Viral replication was followed by measuring the RT activity in the culture supernatant. (C) Schematic of the sequence analysis of the revertant viruses. The cell pellet from (B) was resuspended as described in section 2.6.4. The integrated provirus was amplified by PCR and analysed by sequencing. The sites of the original mutations (X) and second site reversions (X) are indicated. The sequencing data for SM5 were obtained from two independent virus cultures. TAR, *trans*-activation responsive; PA, PolyA; PBS, primer binding site; Ψ, Psi; SD, splice donor; MA, matrix; CA, capsid; NC, nucleocapsid; p6, late domain; P1 and P2, spacer peptides 1 and 2. Schematic not to scale.

In the case of Pal and SM8, it seemed clear that the viruses had mutated as the replication kinetics of the passaged viruses were faster than those of the original viruses (compare Figure 4.8A and B). The situation was not as clear for SM5 as the passaged virus did not start to replicate properly until day 15 post-infection on two independent occasions. This may in part be due to the fact that the virus was passaged before it had reached the peak of infection. Nevertheless, there was still an improvement in the replication kinetic as compared to the original virus.

To confirm whether these mutants had reverted, the integrated proviruses were amplified by PCR using specific primers for the 5′ leader, the U3 promoter region and the *gag* ORF and analysed by sequencing as described in section 2.6.4. Sequencing of Pal and SM8 was performed on a single culture of passaged virus, while two independent cultures of SM5 were used. However, it is important to note that the method used does not allow linking the sequencing data to individual viral clones. A schematic of the sequence analysis results is presented in Figure 4.8C. In all cases, the original mutation had been retained, implying that a second site mutation was responsible for the amelioration in viral replication. Sequencing of the U3 promoter region did not yield any results. However, when the *gag* ORF was sequenced, several mutations were found. Most notably, a T70I mutation in MA occurred in two independent occasions, in SM5 and SM8, and Pal showed a similarly located mutation at amino acid 68 (V68E). Other second site mutations were observed in CA (T98A and V26I) and the late domain p6 (E35K). Surprisingly, no mutations occurred in NC, which corresponds to the RNA-binding domain of the Gag polyprotein.

These compensatory mutations may have an effect at the nucleic acid or at the amino acid level. At the nucleic acid level, one might envisage that the reversion could restore or disrupt a long range RNA interaction or a protein-RNA interaction. However, this

region of MA has not been previously linked with such protein-RNA or RNA-RNA interactions.

Regarding the modifications introduced in the protein sequence, while the V26I mutation in CA may not result in a change in the charge or function of the protein, as both amino acids are aliphatic and hydrophobic, other substitutions could have an impact on the function or charge of the protein.

The T70I mutation, a replacement of a threonine (slightly polar, uncharged and hydrophobic) with an isoleucine (aliphatic and hydrophobic), represents a relatively neutral mutation as both amino acids are hydrophobic and carbon-beta branched, a feature that results in more restriction in the conformation of the main chain (Betts and Russell, 2003). On the other hand, the V68E mutation of an aliphatic and hydrophobic valine substituted with a glutamate (negatively charged, acidic and hydrophilic) did not correspond to a favoured substitution. Interestingly, these two mutations occurred in a region of MA that is possibly quite restricted regarding the conformation of the main chain due to a sequence rich in isoleucine, threonine, valine and proline (60-ILTVLDPMVPT-70).

The T98A mutation in CA, corresponding to the replacement of a threonine with an alanine (hydrophobic), represents a neutral substitution. Finally, in the E35K substitution in p6, a negatively charged acidic amino acid (glutamate) is replaced with a positively charged basic residue (lysine), possibly triggering important changes in the function of the protein even though this substitution is neutral in intracellular proteins and favoured in membrane-associated proteins (Betts and Russell, 2003).

The method used in this study did not allow the analysis of individual clones and it is not possible to determine whether the second site mutations observed were sufficient to rescue viral replication on their own or whether a combination of these mutations was

required. Previous examination of SIV revertants have suggested that sets of several compensatory mutations were required for a full rescue of infectivity, yet, each individual reversion was found to improve the replication of the mutated virus (Guan et al., 2001a). The fact that the SM8 revertant showed only one compensatory mutation (MA T70I) would suggest that it is sufficient to rescue, at least in part, viral replication. However, the presence of further mutations in the genome of these revertant viruses could not be determined due to technical difficulties during the PCR-based amplification and the sequencing steps and it is possible that more second site mutations are present on the genome of these revertant viruses. Therefore, although studies of HIV-1 and SIV mutants impaired for RNA packaging and viral replication have reported compensatory mutations in Gag and the leader region only (Guan et al., 2001a; Liang et al., 1998; Liang et al., 1999b; Liang et al., 2000; Rong et al., 2001; Rong et al., 2003; Whitney et al., 2002), it is impossible to exclude the possibility that other reversions have occurred and played a role in restoring viral replication.

4.4 Discussion

HIV-2 infectivity correlates with genome dimerisation and the presence of the 392-395 GGAG motif

Using a series of mutations, the role of the SL-1 sequence and structure in genome dimerisation, packaging and viral replication was investigated. The results obtained demonstrated that a correlation existed between the ability of the virus to dimerise its genome and its replicative fitness. In particular, it was observed that a purine-rich motif at position 392-395 of the HIV-2_{ROD} genome was absolutely critical for viral infectivity. Removal of that GGAG motif systematically resulted in the production of non-

infectious virions that could not replicate in T-cells and contained mostly monomeric genomes. This result was in agreement with another study that reported a GGRG motif (where R is a purine) at the same position as being important for HIV-2 replication (Baig et al., 2009). Consistent with its essential role in the virus life cycle, that purine-rich motif is well conserved in all HIV-2 and SIV_{SM} subtypes as well as in related SIV species (HIV sequence compendium 2008).

Interestingly, mutation of a second GGAG motif at position 439-442 also had some effect on genome dimerisation and resulted in a delayed replication in T-cells. This suggested that this motif might play a role in viral replication. The HIV-1 Gag polyprotein has been shown to bind purine-rich motifs, which constitute the packaging signal (Berkowitz et al., 1993; Clever et al., 1995). Hence, one might envisage that these motifs could act as a Gag binding site in HIV-2 as well. Consistent with this hypothesis, mutation of the second GGAG motif resulted in a twofold reduction in RNA packaging. It is also interesting to note that whilst most Psi/SL-1 mutants showed some reduction in RNA packaging, the SM4 mutant, which contained an additional GGAG sequence, was the only one to display a packaging efficiency of 90 %, apart from the SM1 mutation of the terminal SL-1 palindrome (see Chapter 3).

The binding affinity of Gag to the mutated RNAs was assessed *in vitro* using EMSA and RNA-protein GST-pull down. Unfortunately, due a lack of sensitivity and reproducibility, these experiments were inconclusive.

Contrary to what was observed with the Psi mutants in Chapter 3, encapsidation did not absolutely correlate with dimerisation. This difference could be explained by the fact that a much greater number of mutations were analysed in the second set of experiments. Furthermore, the Psi mutations DM and Pal both corresponded to rather

large deletions (28 and 10 nt respectively) and could affect dimerisation and encapsidation simultaneously due to the overlap of the RNA elements required for each process. Specifically, these two mutations of the packaging signal also removed the GGAG motif at position 392-395 that was important for genomic RNA dimerisation; hence, one would expect both their dimerisation and packaging abilities to be affected. It would be interesting to test non-radioactive-based methods to assess the packaging efficiencies of the mutant viruses to confirm the results of the RPAs as the use of radioactive methods presents the major problem of the decay of the radioisotope. The use of radioactivity in RNA-protein binding experiments also posed difficulties. Quantification of the RNA input in the EMSA using TCA precipitation did not prove the most reproducible method and it was difficult to perform each experiment on an equivalent amount of RNA, rendering the results impossible to interpret. Similarly, measure of the RNA bound to GST-Gag in the GST pull-down experiment suffered from a lack of reproducibility due to variation in both the RNA input and the output reading of the experiment by scintillation counting. Methods based on real-time RT-PCR are ideal candidates to replace the use of radioactivity in the packaging assays, as the concentration of the RNA input could be measured precisely using a Nanodrop spectrophotometer. Besides, real-time RT-PCR is an extremely sensitive method, allowing detection of minimal amounts of RNA.

However, despite apparent discrepancies between the results obtained with the two sets of mutants (presented in Chapters 3 and 4) regarding the connection between genome dimerisation and packaging, a nonreciprocal relationship is observed between the two processes. When the set of data is analysed as a whole (Table 4), one can note that all mutants that failed to efficiently form RNA dimers also showed a packaging defect, yet

the reverse is not true. This suggests that genome dimerisation is important for encapsidation and it implies that RNA dimer formation occurs prior to packaging.

Recent work by the Lodmell group has shown that a GGRG sequence in the HIV-2 Psi region was important for packaging as well as replication (Baig et al., 2009). Our data did not corroborate this finding and mutants that contained an intact GGAG motif failed to package their genomes to wild type level (e.g. SM6, SM7 and SM9). In their study, Baig and co-authors used an *in vivo* SELEX (selected evolution of ligands by exponential enrichment) method, in which randomised 5′ *pal* sequences were introduced into a pool of genomes and selected through serial passages of the viruses. Sequencing of the surviving viruses revealed a consensus towards a GGRGN sequence at position 392-396, suggesting that a purine-rich motif was indeed critical for viral replication. Analysis of the packaging ability of two pyrimidine-rich mutants, TTB16 (5′-ACCUG-3′) and the SM2 mutant described in this study (5′-CCUCU-3′) and previously published (L'Hernault et al., 2007), as well as a purine-rich mutant, TTB78 (5′-GGGGU-3′), showed that only the pyrimidine-rich mutants had a reduction in packaging, two- to threefold compare to the wild type virus, leading to the conclusion that a purine-rich motif was critical for encapsidation. In the present study, six mutants were analysed for their ability to dimerise and encapsidate their genomes and replicate in T-cells and a careful comparison between the sequences and structural elements present in each mutant and the experimental data obtained for each of them was carried out. Extending this comparative analysis to the Psi mutants presented in Chapter 3 (see Table 4), it was confirmed that the ability to form RNA dimers correlated with the packaging efficiency and the replicative fitness of HIV-2 and that the presence of the GGAG motif at position 392-395 was essential for HIV-2 infectivity. In our study, genome encapsidation did

not, however, depend upon the presence of this purine-rich motif. Further analysis will be required to clarify the requirement for a purine-rich motif - such as GGRG - in HIV-2 RNA packaging.

Table 4: Comparative analysis of the Psi and SL-1 mutants.

Virus	Dimerisation *in virio* [a]	Packaging [b]	Replication [c]	*pal*	DIS	Stem B[f]	392-395 GGAG	439-442 GGAG
WT	+++	1	+	+	+	+	+	+
SM1	+++	0.9	+	+	-	+	+	+
SM7	++	0.5	+	+	+	+	+	+
SM9	++	0.5	+	-	+	+[g]	+	-
SM4	+	0.9	+	-	+	-	+	+
SM6	+	0.5	+[d]	+	+	-	+	-
DM	-	0.4	-	-	+	-	-	+
Pal	-	0.4	-	-	+	-	-	+
SM2	-	0.5	-	-	+	+	-	+
SM3	-	0.4	-	-	-	+	-	+
SM5	-	0.6	-	+[e]	+	-	-	+
SM8	-	0.4	-	-	+	-	-	-

[a], genomic RNA dimerisation as measured by native northern blot.
[b], packaging efficiency determined by RPA. Wild type (WT) is set as 1.
[c], Replication in PM1 T-cells, followed by measuring the RT activity in the culture supernatant. Presence (+) or absence (-) of detectable RT activity over a 20 days monitoring period.
[d], a 2-days delay in replication was observed prior to wild type-like replication.
[e], alternative palindromic sequence in place of wild type *pal*.
[f], base-pairing between residues 397-401 and 439-443.
[g], alternative sequences restoring the base-pairing forming in stem B.

Purine-rich sequences are involved in a wide range of cellular processes

Purine-rich motifs can be found in a variety of organisms and have been proposed to play a role in post-transcriptional control, mRNA splicing, translation, protein-RNA interactions and RNA packaging.

For example, in the bacteriophage T4, GGAG motifs are involved in post-transcriptional control as early mRNAs are cleaved by the endoribonuclease RegB, mostly at GGAG sites located in intergenic regions - including Shine-Dalgarno sequences - resulting in their inactivation (Jayasena et al., 1996; Sanson and Uzan, 1995).

In humans, members of the serine/arginine-rich (SR) family of proteins such as splicing factor 2 (SF2/ASF) and SRp30, which are involved in pre-mRNA splicing, mRNA export and translation, have been shown to bind to purine-rich sequences - mainly located in exons - resulting in the enhancement of splicing (Sanford et al., 2008; Tian and Kole, 2001). Interestingly, the N-terminal region of HIV-2 Gag has been shown to specifically interact with a serine-threonine kinase from the SR family of splicing factors, PRP4, resulting in the inhibition of SF2/ASF phosphorylation (Bennett et al., 2004). However, this interaction, which was demonstrated by yeast-two-hybrid, did not appear to involve an RNA intermediate.

Another splicing factor, SFRS1, was shown to recognise a purine-rich octamer (Sanford et al., 2009), similar to the sequence located between SL-1 and the SD stem-loop in HIV-2.

In higher eukaryotes, proteins are translocated through the RER membrane during their synthesis where they undergo a range of post-translational modifications. To achieve this, the signal recognition particle SRP, a cytoplasmic ribonucleoprotein complex, directs the ribosome to the RER membrane and facilitates the insertion of the nascent polypeptide into the membrane (reviewed in Lutcke, 1995). Assembly of the human SRP has been shown to include the specific binding of the SRP19 protein on the SRP RNA, at a site that comprises a central GGAG loop (Sakamoto et al., 2002).

In vivo footprinting carried out in infected T-cells also showed that several purine-rich sites in the HIV-1 promoter region, sharing the common sequence 5′-GGAGAGA-3′, were prime protein binding sites (Demarchi et al., 1992), suggesting that these purine-rich motifs are also involved in DNA-protein interaction and might play a role in HIV-1 mRNA transcription.

Finally, purine-rich motifs have been found to be key players in protein-RNA interactions in several viruses, including HIV-1. Interestingly, it was shown that the coat protein of Cauliflower Mosaic virus interacts with purine-rich sequences on the pre-genomic RNA via its zinc finger motif and basic residues (Guerra-Peraza et al., 2000). This mechanism displays extreme similarities with the NC-RNA interaction observed in retroviruses during genome packaging. The zinc fingers in HIV-1 and MLV NC and the flanking basic amino acid residues in HIV-1 NC have indeed been reported to be required for this interaction (Aldovini and Young, 1990; Berkowitz and Goff, 1994; Dannull et al., 1994; De Guzman et al., 1998; Gorelick et al., 1988). The NC domain of the HIV-1 Gag polyprotein binds to a sequence in the 5′ leader RNA located close to the Gag start codon (SL-3) which corresponds to the main packaging determinant and contains a GGAG motif exposed in a loop (Aldovini and Young, 1990; Bacharach and Goff, 1998; Berkowitz and Goff, 1994; De Guzman et al., 1998; Lever et al., 1989; Luban and Goff, 1994). Interestingly, a GA-rich sequence located immediately downstream of SL-3 was also found to be involved in HIV-1 RNA packaging and dimerisation (Russell et al., 2003a) and *in vitro* footprinting experiments confirmed that purine-rich sequences were amongst the preferred binding sites for HIV-1 Gag and NC (Damgaard et al., 1998).

Possible roles of the HIV-2$_{ROD}$ 392-395 GGAG motif

The position of the 392-395 GGAG motif in the 5′ untranslated region of the HIV-2 genome and the similarities with other retroviruses, including HIV-1, suggest a role in protein-RNA binding and RNA packaging. Indeed, a function in regulation of mRNA splicing seems less likely as the purine-rich splicing regulatory elements described above appear to be primarily located in exons. Besides, the binding site of the splicing enhancer SFRS1 shares strong homology with the GA-rich stem-loop located between SL-1 and the SD stem-loop in the HIV-2 leader RNA, but not with SL-1 itself.

With regards to a role in mRNA translation, there is no evidence so far to support this hypothesis, as the translation of the GGAG mutants appeared normal. Nonetheless, this possibility cannot be completely excluded since the rate of protein synthesis and processing of the mutants has not been measured. One might envisage that similarly to what is observed in HIV-1, where the Gag protein modulates its own translation partly through binding of NC to the RNA leader (Anderson and Lever, 2006), the HIV-2 Gag protein also has some self-regulatory function involving sequences in the RNA leader. In HIV-1, the translation of Gag is stimulated when the protein is present at low concentrations and inhibited at high concentrations, probably corresponding to a switch between the need for the structural protein synthesis and the need for packaging and assembly. The matrix domain appeared to stimulate Gag translation - independently of RNA binding - while binding of the NC domain to the packaging signal seemed to cause the inhibition of translation (Anderson and Lever, 2006). In HIV-2, only the inhibitory effect is seen and it appears that the matrix domain does not stimulate Gag translation (Emma Anderson, personal communication). This may be due to the inability of HIV-2 MA to efficiently bind to the eukaryotic initiation factor 5B or eIF5B (Andrew Lever, personal communication) as one hypothesis to explain the stimulatory

effect of HIV-1 MA on Gag translation is that HIV-1 MA, which was shown to interact with eIF5B (Andrew Lever, personal communication; Wilson et al., 1999), could act as a guanine exchange factor for eIF5B (Emma Anderson, personal communication). If HIV-2 Gag also modulates its own translation through the binding of NC to the 5′ leader RNA, mutation of a Nucleocapsid binding site would result in a shift in the balance between translation and packaging. In particular, considering that HIV-2 packages its genome in a co-translational manner and that the limited availability of the Gag polyprotein ensures the specificity of packaging, one could envisage that prevention of the NC-mediated inhibition of Gag translation would affect HIV-2 genome encapsidation. This might provide an explanation for the reduced packaging and delayed replication of the SM6 virus that bears a mutation at a second GGAG motif at position 439-442, another putative NC binding site. It would be interesting to test this hypothesis by measuring the effect of HIV-2 NC on Gag translation in the context of the SL-1 mutants described in this study.

Potential implications of the compensatory mutations in the gag coding sequence

While all mutations that removed the 392-395 GGAG motif resulted in an absence of detectable replication in PM1 cells over the course of 20 days, prolonged culture of these non-infectious viruses resulted in the emergence of revertants for SM5, SM8 and Pal. All revertants retained the original mutations, suggesting that the rescue of viral replication was triggered by second site compensatory mutations. The promoter region, 5′ leader and *gag* ORF of Pal, SM5 and SM8 revertants were sequenced and several mutations were identified in the *gag* coding region. Each virus carried a mutation in MA at position 68 or 70. Pal displayed a V68E substitution and SM5 and SM8 both had a T70I substitution. Interestingly, the MA T70I mutation has been observed previously in

a SIV_{MAC239} packaging mutant bearing a deletion at position 398-418 (termed SD2) in combination with a mutation in RT (M184V) that delays reversion (Whitney et al., 2002). Residues 398-418 of SIV_{MAC239}, which are almost 100 % conserved with nucleotides 397-417 of HIV-2_{ROD} and predicted to fold into a structure similar to SL-1 in HIV-2_{ROD}, appeared to be required for viral replication in SIV (Guan et al., 2000). The T70I reversion conferred a replicative advantage to the SD2-M184V virus but did not restore replication to wild type level (Whitney et al., 2002), similarly to what can be seen with the SM5 and SM8 revertants. The impact of this reversion on the replication of the SD2 mutant in the absence of the M184V mutation was unfortunately not assessed.

The T70I compensatory mutation was also identified in HIV-1 MA mutants bearing substitutions at residues 4 and 6 and presenting defects in particle assembly and release. In HIV-1, the T70I reversion was found to induce a modest increase in the rate of Gag p55 processing, but had no effect on the kinetics of replication (Ono et al., 1997).

The MA protein has been implicated in RNA dimerisation in RSV as mutations that altered the localisation of MA were found to disrupt dimer formation (Garbitt et al., 2001). It has also been proposed that HIV-1 MA is involved in the proteolysis of the Gag precursor during assembly (Lee and Linial, 1994; Ono et al., 1997). Furthermore, Gag multimerisation, which is required for the assembly of retroviral particle and is enhanced by the presence of RNA, has been shown to require the capsid-dimer interface and the basic residues in MA in the absence of NC (Burniston et al., 1999). This suggests that protein-protein and protein-RNA interactions may be involved in the assembly process and that MA might be involved in some of these protein-RNA interactions via its basic residues. In addition, HIV-1 MA has been shown to play a role in the targeting of the Gag polyprotein to the plasma membrane as part of the assembly

process (Facke et al., 1993; Ono and Freed, 1999; Ono et al., 2000) and in the incorporation of the Env glycoprotein into virions (Dorfman et al., 1994b; Freed and Martin, 1996; Ono et al., 1997). HIV-1 MA, but not HIV-2 MA, has also been shown to stimulate Gag translation when Gag is present at low concentration, possibly via interaction of the N-terminus of matrix with eIF5B (Anderson and Lever, 2006; Lever, personal communication). The region between residues 32 and 57 of HIV-1 MA was shown to be required for the interaction with eIF5B (Lever, personal communication). Finally, these mutations are located immediately downstream of the second Gag AUG, proposed to promote the translation of the p50 isoform of Gag (Herbreteau et al., 2005), and could affect the production of the p50 isoform and/or the ratio of the three Gag isoforms produced. Unfortunately, the Gag protein synthesis of the revertant viruses was not assessed.

Interestingly, at the nucleic acid level, part of the *matrix* coding region also appears to be involved in a long-range pseudoknot RNA-RNA interaction with sequences downstream of the PolyA signal (Paillart et al., 2002). Sequences required for this interaction were found to be conserved in HIV-2 and SIV isolates, supporting its functional importance; yet, the exact role of this interaction has not been elucidated. It was suggested that it could act as a repressor of the polyadenylation signal, regulate steps of the life cycle such as Gag translation and/or facilitate the selection of the genomic RNA during packaging.

In addition to the reversions in MA, mutations in CA (V26I and T98A) and p6 (E35K) were observed for the HIV-2 SM5 and Pal mutants.

In a previous study of the SIV_{MAC239} SD2 mutant, which showed the same T70I reversion in MA as SM5 and SM8, two sets of three compensatory mutations were found, including mutations in CA and p6 (Guan et al., 2001a). In each case, an A423G

substitution in the terminal palindrome of SL-1 was observed. This mutation disrupted the putative DIS and would have prevented self-complementarity (GGUACC to GGUGCC). In addition to this mutation, two mutations in CA (K197R) and p6 (E49K) were identified when the virus was grown in C8166 cells. Using CEMx174 cells, in which the replication kinetic of the viruses was much faster, two mutations in NC (E18G and Q31K) were found. For each set, the three mutations were required to restore full replication, even though each individual reversion on its own was able to provide a replicative advantage. Similar studies of HIV-1 SL-1 and SL-3 deletions mutants impaired for RNA dimerisation and packaging also found compensatory mutations in Gag. The SL-1 mutant reverted in MA (V35I), CA (I91T), p2 (T12I) and NC (T24I), while the SL-3 mutant mutated p2 (A11V) and NC (I12V) (Liang et al., 1999b; Rong et al., 2003). Importantly, the T24I mutation in NC appeared first and contributed most to the rescue of viral replication. While it seems probable that mutations in NC would rescue the replication of a mutant virus with a defect in viral RNA packaging, as NC plays a major role in RNA selection during the encapsidation process, it is less clear what effects the compensatory mutations in CA and p6 had.

Capsid is the major component in particle assembly and the presence of RNA has been shown to enhance VLPs assembly *in vitro* (Campbell and Vogt, 1995). Hence, it is possible that the reduction in the level of dimeric RNA packaged affected the stability of the viral capsid and that these CA compensatory mutations restored it. This hypothesis would depend on the level of dimeric RNA packaged in the revertant virus as one might expect packaging, and possibly dimerisation, to be at least partially restored in those viruses; otherwise, another explanation for the reversion in CA is required. V26 is located in the alpha helix 1 of CA NTD. Interestingly, residues in helix 1 of HIV-1 CA have also been linked with genome dimerisation as it was shown that

mutations that disrupt alpha helix 1 but leave other region of CA intact - in particular the C-terminal domain - resulted in a dimerisation defect similar to that observed in protease-deficient viruses (Kafaie et al., 2009). Furthermore, helix 1 to 3 of HIV-1 CA have been found to play a role in mature capsid formation (von Schwedler et al., 2003a) and the V26I mutation might affect particle assembly or structure. In fact, the V26I mutation has been observed previously in HIV-1 in the context of a W23F CA mutant displaying a replication and assembly defect (Tang et al., 2007). In that study, it was reported that the NTD of CA and in particular residues involved in alpha helix 1 and 2 contribute to proper core assembly and that the V26I reversion of the W23F mutant partially restored the conformation of CA.

The E35K mutation in p6 is located within the Vpx binding site (Pancio and Ratner, 1998) and near the PTAPP motif which was shown to interact with Tsg101, an inactive E2 ubiquitin ligase homologue, triggering an increased ubiquitination of HIV-2 Gag (Myers and Allen, 2002). As monoubiquitination of Gag is important for its targeting to multivesicular bodies, the introduction of an extra lysine, which could be ubiquitinated, may alter the trafficking pattern of Gag.

The method used in this study did not allow us to determine whether either compensatory mutation alone was able to restore replication or if sets of reversions were necessary. However, the fact that the SM8 revertant carried only the MA T70I mutation and that mutation in MA were found in all three revertants suggested that mutations in this region of Gag conferred a significant replicative advantage to the virus.

Whilst the exact function of these compensatory mutations in MA, CA and p6 has not been elucidated, the aforementioned studies in HIV-1 and SIV seem to point towards a role either in RNA-protein or protein-protein interaction, possibly to help restore the

packaging process, stabilise the RNA dimer or enhance/rescue retroviral assembly should the mutations introduced in SL-1 have impaired capsid formation.

It is important to note that the genome of the revertant viruses was not entirely sequenced due to technical difficulties during the amplification and sequencing processes. As a result, despite several studies of HIV-1 and SIV packaging mutants reporting the presence of compensatory mutations in the 5′ leader and *gag* coding region only (Guan et al., 2001a; Liang et al., 1999b; Rong et al., 2003; Whitney et al., 2002), the possibility of more compensatory mutations elsewhere on the genome cannot be ruled out and it is impossible to draw any definitive conclusions regarding the reversions found in the *gag* coding sequence. To do so, a complete sequencing analysis of the integrated provirus as well as the cloning of the candidate reversions into the molecular clones of the mutant viruses is necessary.

5 Role of RNA dimer formation in particle maturation

5.1 Introduction

Nucleic acids and RNA in particular have been suggested to play an important role in the correct assembly of retroviral particles. Although early studies of HIV-1 packaging mutants reported normal protein and particle production, as measured by the level of CA (p24) in the cells and culture supernatant (Aldovini and Young, 1990; Clavel and Orenstein, 1990; Gorelick et al., 1993; Gorelick et al., 1990; Lever et al., 1989; Luban and Goff, 1994), differences in particle morphology were observed. In some reports, HIV-1 Psi mutants appeared morphologically similar to wild type viruses (Gorelick et al., 1990; Lever et al., 1989), whereas in others, particles with an aberrant morphology - empty or less dense particles - were described (Aldovini and Young, 1990; Clavel and Orenstein, 1990), suggesting that normal genomic RNA packaging may be required for the orderly arrangement of the nucleoid components. A subsequent study in RSV and HIV-1 showed that the efficiency of *in vitro* assembly of purified CA-NC and CA-NC-p6 into VLPs was stimulated by the addition of nucleic acid (Campbell and Vogt, 1995). Similar observations were made using an HIV-1 Gag protein lacking only the p6 domain (Campbell and Rein, 1999). In both studies, the multimerisation of the Gag polyprotein into units larger than dimers was greatly enhanced by the presence of nucleic acid (Campbell and Rein, 1999; Campbell and Vogt, 1995), which support the hypothesis that particles assemble through the formation of an RNA:protein complex that serves as a nucleation core for the expansion of the viral capsid in an RNA-

dependent manner. Further evidence that RNA was acting as a bridge in Gag-Gag interaction was provided by analysis of several Gag deletion mutants showing that the nucleic acid binding domain of Gag, NC, was required for the interaction and that RNAse treatment disrupted Gag multimerisation *in vitro* (Burniston et al., 1999). Accordingly, HIV-1 NC mutants defective for specific and non-specific RNA binding showed impaired Gag-Gag interaction, suggesting that RNA binding is indeed required for Gag multimerisation (Cimarelli et al., 2000). Interestingly, this implication of RNA in Gag multimerisation and particle assembly is in accordance with the 'fullerene' model proposed for HIV-1 capsid assembly (Ganser et al., 1999; Li et al., 2000; Nermut et al., 1994).

The addition of RNA was also found to increase the efficiency of CA-NC particles assembly, which otherwise required very high salt concentration (Ganser et al., 1999). However, it is interesting to note that this stimulatory effect was not found to be specific of the viral RNA and that several nucleic acids such as yeast tRNA, bacterial rRNAs, tobacco mosaic virus RNA and DNA oligonucleotides were also found to enhance the assembly of CA-NC *in vitro* (Campbell and Rein, 1999; Ganser et al., 1999). Finally, the presence of nucleic acid in *in vitro*-assembled HIV-1 VLPs was confirmed by immunogold staining (Fuller et al., 1997).

Interestingly, a study of MLV reported that particles that lacked genomic RNA due to a mutation in the packaging signal contained cellular RNA in place of viral RNA and that RNase treatment disrupted the virus cores, suggesting that RNA was a structural element of MLV particles (Muriaux et al., 2001). In addition, RNA was proposed to be required for the proteolytic cleavage of the HIV-1 NC (p15) precursor (Sheng and Erickson-Viitanen, 1994), suggesting that binding of the NC protein to RNA might play a role in virion maturation.

In HIV-1, deletions around the DIS were reported to cause a delay in the processing of the p2 peptide from the CA-p2 precursor (Liang et al., 1999a). Since the p2 peptide has been demonstrated to be involved in the sequential proteolytic processing of the Gag protein and the production of fully infectious particles (Pettit et al., 1994), this result implied that sequences important for RNA dimer formation play a role in protein processing and consequently in particle maturation. However, it was not possible to conclude whether the formation of an HIV-1 RNA dimer was important as the dimerisation abilities of these particular DIS-mutated viruses had not been established (Liang et al., 1999a). Nonetheless, further evidence that the dimerisation or the packaging of the genomic RNA contributed to correct particle assembly came from a study in SIV, where mutations in the SL-1 region that resulted in a reduction of encapsidation and dimerisation led to aberrant particle morphology (Whitney and Wainberg, 2006).

Whilst several reports have proposed a role for the RNA - possibly as a dimer - in particle assembly and morphogenesis, the HIV-1 and MuLV RNA dimers have been shown to undergo a 'maturation' process dependent on proteolytic cleavage by the viral protease (Fu et al., 1994; Fu and Rein, 1993) and therefore likely to occur simultaneously to the maturation of the viral particle. Gag cleavage alone is however not sufficient to promote the maturation of the dimer (Fu et al., 2006) and other factors might be involved. Interestingly, the proteolytic processing of the CA/p2 cleavage site, the ratio of Gag and Gag-Pol polyprotein synthesised and certain residues in the spacer peptide p1 have also been implicated in the stability and maturation of the HIV-1 RNA dimer (Hill et al., 2002; Shehu-Xhilaga et al., 2001a; Shehu-Xhilaga et al., 2001b).

In this chapter, we explored the link between the encapsidation and dimerisation of the HIV-2 genome and the correct assembly of the virus particle. Wild type and DM mutant

viruses were analysed by electron microscopy using negative staining. Micrographs of the DM virus revealed a decrease in the number of mature particles compared to the wild type virus. Yet, no increase in the number of immature particles was observed. Instead, an increased number of particles containing two or more cores was noted. The data obtained suggested that the DM mutant virus partly failed to produce adequately matured particles and that this maturation defect might in part account for the reduced infectivity observed with this mutant.

5.2 Effect of the DM mutation on particle morphogenesis

The synthesis and processing of the Gag polyprotein in the DM mutant has been assessed previously using western blotting and metabolic labelling (see Chapter 3 and Griffin et al., 2001). However, these analyses were conducted at a single time point of 48 h post-transfection and are not sufficient to exclude the possibility of a delay in Gag or Gag-Pol processing, which could affect the maturation of the virus and account for the reduction in viral infectivity of this mutant virus.

In order to evaluate the impact of the Psi deletion on particle morphology and determine whether there was any contribution of RNA encapsidation and genome dimerisation to viral assembly in HIV-2, wild type and DM virions - produced by Cos-1 cells transfected with the appropriate envelope-deleted provirus - were purified through 20 % sucrose and analysed using negative stain electron microscopy. As in experiments performed on purified HIV-1 cores (Welker et al., 2000) and *in vitro*-assembled HIV-1 capsids (Campbell and Rein, 1999), samples were stained with 1 % uranyl acetate prior to observation by electron microscopy. Electron micrographs were examined by two independent observers and approximately two hundred particles were identified for each

virus. Representative particles containing immature, mature and multiple cores are shown in Figure 5.1. Immature particles displayed a peripheral ring, likely composed of the structural polyprotein Gag orderly arranged below the lipid membrane as observed in HIV-1 (Fuller et al., 1997; Gelderblom, 1991; Hockley et al., 1988; Wilk et al., 2001). HIV-2 mature particles appeared to contain a cone-shape core, as previously noted for HIV and SIV (Briggs et al., 2003; Gelderblom, 1991; Grief et al., 1989; Hockley et al., 1988; Kewalramani and Emerman, 1996; Palmer et al., 1988; Welker et al., 2000) and in accordance with the 'fullerene'-model (Ganser et al., 1999; Li et al., 2000; Nermut et al., 1994). Finally, some particles containing two cores were noticed. In most cases, the two cores were of conical shape, though some tubular cores were also detected (Figure 5.1). These type of particles, which may have resulted from the assembly of two cores near to each other and their subsequent budding in a single envelope, have been described previously (Briggs et al., 2003; Gelderblom, 1991) and were found to represent a third of mature HIV-1 particles (Briggs et al., 2003).

Approximately two hundred wild type and DM particles were observed and classified according to their morphology and the results are presented in Figure 5.2. Interestingly, the preparation of purified DM virus was found to contain significantly fewer mature particles compared to the wild type virus preparation (Figure 5.2A). However, this reduced number of mature particles for the DM virus did not correlate with an increase in the number of immature virions (Figure 5.2A), suggesting that the particle maturation process may not be blocked at an early stage of the proteolytic processing. In contrast, there was a larger number of aberrant particles containing two cores in the DM sample (Figure 5.2A, multiple cores). The remaining proportion of the particles appeared to contain a core located off-centre (Figure 5.2A, eccentric cores). These particles may have been mature, as they contained the cone-shaped core typical of mature particles,

yet, due to the eccentric position of the core inside the particle, these were classified and counted as a separate category. No difference was observed between the wild type and mutant virus preparation for these eccentric-core particles.

Finally, it was also noted that the average diameter of the mutant virions was larger than that of the wild type virions (Figure 5.2B). This did not appear to result from the increased number of particles containing multiple cores, as the average diameter of the DM virions (186.7 ± 31.8 nm) was the same as the average diameter of the multiple core DM particles only (185.9 ± 22.6 nm; compare Figure 5.2B and C). It is interesting to note that an increase was observed between the average size of the total wild type virions (157.7 ± 38.5 nm) and the wild type multiple core particles only (169.3 ± 21.6 nm), which could reflect the presence of the second core inside these particles; yet, this difference was not statistically significant.

The data presented here showed that a mutation in the HIV-2 packaging signal that also affected genome dimerisation resulted in a reduction in the number of mature particles and an abnormal proportion of particles containing more than one core and may in part explain the loss of infectivity associated with this mutation.

5.3 Discussion

Negative staining electron microscopy was used to analyse whether the HIV-2 packaging mutant displayed a morphological or maturation defect which could explain the reduction in infectivity observed. Due to safety issues, the experiments were carried out using envelope-deleted virions.

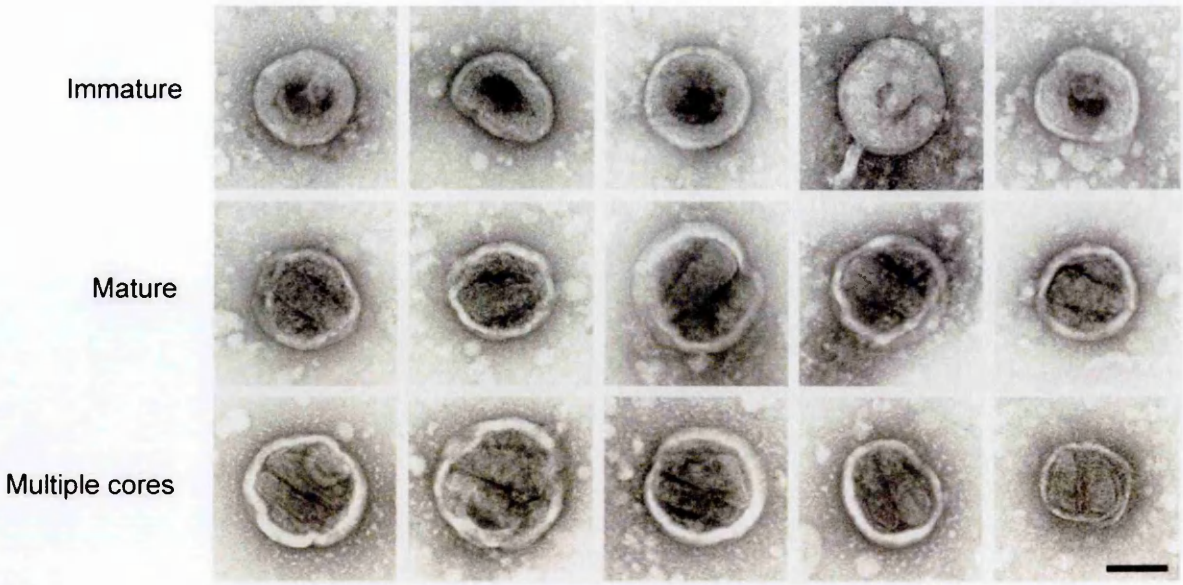


Figure 5.1: Representative electron micrographs of HIV-2 particles.
Envelope-deleted wild type and DM mutant virions were produced by transfection of Cos-1 cells with the appropriate molecular clones. At 48 h post-transfection, virions were harvested and purified through a 20 % sucrose cushion. Virions were then washed and stained as described in section 2.6.5 prior to visualisation by electron microscopy. Representative micrographs of three independent microscopy sessions using two independent virus preparations are shown. Five examples each of particles denominated as immature, mature and containing multiple cores are shown. Magnification: 20,000 ×; scale bar: 100 nm.

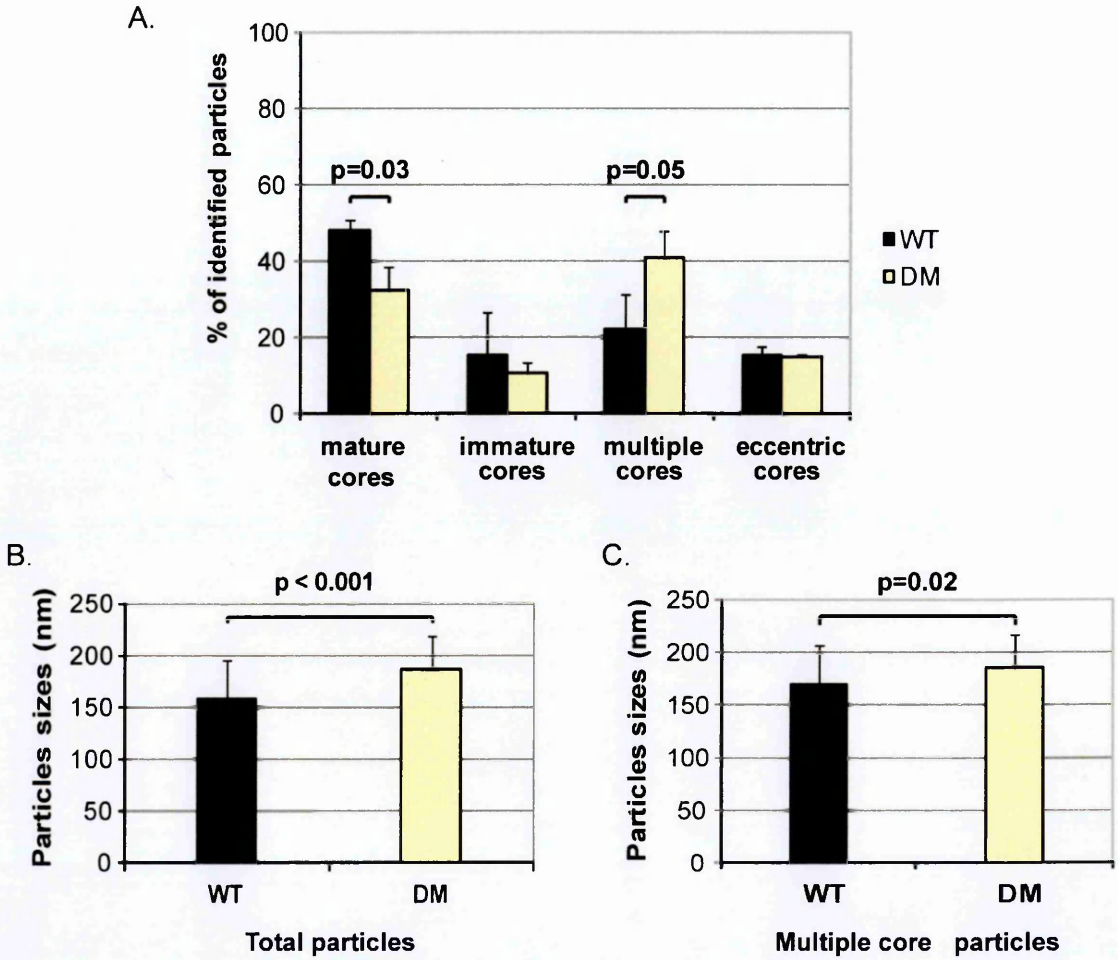


Figure 5.2: Effect of the Psi deletion on HIV-2 particle morphology.

(A) Quantification of wild type (WT) and DM mutant HIV-2 particles observed by negative staining electron microscopy. Immature, mature and multiple core particles are as described in Figure 5.1. Particles containing a single core in an off-centre position (eccentric cores) were also counted. Micrographs of three independent microscopy sessions using two independent virus preparations were analysed by two individuals not knowing which micrograph corresponded to which virus preparation. Average and sd for 193 WT and 199 DM particles is shown. Statistical significance measured by a two-tailed unpaired Student t test. The size of all particles (B) or those containing two cores (C) was measured by a single observer for the WT and DM mutant virus. Average and sd are shown. Statistical significance measured by a two-tailed unpaired Student t test.

To ensure an objective, non-biased analysis, the electron micrographs were analysed by two observers who did not know which micrographs corresponded to which virus. Electron micrographs of the packaging mutant (DM) virus showed a modest, yet significant reduction in the number of mature particles compared to the wild type virus. However, the proportion of immature particles in the mutant virus sample was similar to that of the wild type sample, indicating that the maturation of the DM virus was not blocked at an early stage of the proteolytic processing. In addition, protein processing of the DM mutant appeared normal when analysed at the same time point of 48 h post-transfection that was used in this study (Griffin et al., 2001). Nonetheless, a delay in protein processing cannot be excluded as the analysis was only performed at a single time point. Interestingly, a recent analysis of the SIV_{MAC} SL-1 region showed that deletions upstream of the SD, which affected RNA packaging, also had an effect on protein processing and particle morphogenesis (Whitney and Wainberg, 2006). With this in mind, it would be interesting to analyse the protein processing of the mutant virus at several time points and determine whether the processing of the Gag and Gag-Pol polyprotein is affected. Metabolic labelling of the neo-synthesised proteins at several time points would help determine whether the synthesis and subsequent cleavage, as well as the ratio, of the Gag and Gag-Pol polyproteins were correct. Due to a lack of time, this experiment could not be conducted.

As previously observed with HIV-1 (Briggs et al., 2003), a significant proportion (20 %) of wild type HIV-2 particles was found to contain two or more cores. Interestingly, the number of particles containing two cores was increased by twofold in the DM mutant sample compared to the wild type sample. There has been no report regarding the infectivity of these 'two-core' particles. Hence, it is difficult to evaluate whether these represent infectious virions. Yet, the fact that their number was increased in the

replication-incompetent DM virus preparation suggested that they may not be. Therefore, the higher proportion of these multiple core particles might account at least in part for the infectivity defect observed with the DM mutant.

In HIV-1, these particles containing two cores were found to be significantly larger (20 nm) than those containing a single core (Briggs et al., 2003), which suggested that these particles may contain four copies of the genetic material, two in each core. Considering that the DM mutant showed a reduced level of dimeric RNA in the virions, it would be tempting to speculate that each core contained a single copy of genomic RNA, packaged as a monomer. Since the average diameter of the DM particles appeared to be the same as the average diameter of a DM particle containing two cores, it seemed unlikely that the two cores present in a single particle each contained two copies of the genomic RNA. Yet, one could not conclude from this experiment whether each of these two cores contained only a single copy of the genome. In fact, proving this rigorously would be technically challenging, as one would have to design dimer- and monomer-specific probes and label them with gold particles of different sizes to distinguish between monomeric and dimeric RNA within the two cores.

The DM particles appeared to be significantly larger than the wild type particles, by approximately 30 nm on average. Since the DM sample also comprised significantly more particles containing two cores, one could envisage that the higher proportion of multiple core particles accounted for the size difference. However, when the DM particles with two cores were measured, these were found to be of equal diameter compared to the average DM particle, indicating that the increased number of particles containing two cores did not account for the larger diameter of the DM particles. Interestingly though, the average wild type particle containing two cores was found to be larger than the average wild type virion by 11.5 nm. A possible explanation would be

that, in the case of the wild type virus, these particles contained two complete mature cores, each comprising two copies of the genomic RNA. Further experiments would however be required to determine whether this size difference could be justified by the presence of twice the amount of genetic material in these particles.

Although the data obtained here showed that the DM mutant virus showed a maturation defect, even if modest, it remained unclear at which stage of the virion maturation this deficiency occurred. The fact that the number of immature particles did not increase as the number of mature virions decreased indicated that the block was probably not at an early stage of the maturation process. The reduction in the level of RNA packaged in the mutant virus may have resulted in a decrease stability of the virions after assembly at the plasma membrane, as proposed for HIV-1 (Wang and Aldovini, 2002). Indeed, abolition of RNA packaging by mutations of the NC domain did not affect the assembly of Gag particles at the plasma membrane but resulted in a drastic decrease in the level of Gag protein associated with sedimented retroviral particles (Wang and Aldovini, 2002). In particular, the interaction of the NC protein with RNA - viral or not - prevented the disassembly of the retroviral particle during proteolytic cleavage by the viral protease (Wang et al., 2004). In this case, the post-assembly defect or reduced stability of the particles occurring during particle maturation due to inefficient RNA incorporation would not affect the proportion of immature particles released from the cells. Hence, it is possible that the particles produced by the HIV-2 packaging mutant would dissociate rapidly after release due to a lack of RNA packaged in the virions.

As discussed in the previous chapter, long-term culture of non-infectious SL-1 mutants resulted in the emergence of a few revertant viruses bearing compensatory mutations in MA, CA and p6. Interestingly, one possibility was that the second site mutations observed in MA (V68E and T70I) and CA (V26I and T98A) would affect Gag

multimerisation and particle assembly, hence rescuing the viral replication of the mutated viruses. This hypothesis supports the suggestion that mutations in the Psi/SL-1 region result in a defect in particle assembly and/or stability, as observed with the DM mutant. It would have been interesting to analyse the morphology of these mutant virions prior to and after reversion to confirm (i) whether the Psi/SL-1 mutations disrupted capsid formation and (ii) whether the compensatory mutations in CA and/or MA could rescue particle assembly.

The diminution of the number of mature virions correlated with the augmentation of particles that contained two or more cores. These would be predicted to be non-infectious and could have arisen from the targeting and subsequent assembly of two cores at the same location on the plasma membrane, which would then be assembled into a single particle. Whether these 'two-core' particles would be infectious in the case of the wild type virus, where RNA packaging and dimerisation are normal, remains to be elucidated.

The presence of nucleic acid, and in particular RNA, has been shown to considerably enhance the assembly of retroviral VLPs *in vitro* (Campbell and Rein, 1999; Campbell and Vogt, 1995; Fuller et al., 1997; Ganser et al., 1999). In addition, MLV packaging mutants were shown to encapsidate cellular RNA in place of the viral RNA and viral cores were found to be RNAse-sensitive, suggesting that RNA was a structural element of the retroviral particle (Muriaux et al., 2001). The interaction of NC with viral or non-specific RNA was also shown to promote HIV-1 Gag polyprotein mutlimerisation *in vitro* (Burniston et al., 1999; Cimarelli et al., 2000), implying that an RNA:protein complex is required during assembly. Yet, whether RNA dimerisation is required for this process remained unclear. The fact that non-specific RNA could stimulate Gag multimerisation as efficiently as a viral RNA would argue that the RNA does not need

to be dimeric. Several studies have proposed that it is the interaction of the NC domain of HIV-1 Gag with an RNA that is critical for promoting Gag-Gag interaction and maintaining the particle stability during proteolytic cleavage by the protease (Burniston et al., 1999; Cimarelli et al., 2000; Wang and Aldovini, 2002; Wang et al., 2004), again suggesting that the dimerisation efficiency of the viral RNA does not play a role in particle assembly and maturation.

Since it was not possible to discriminate between the packaging and the dimerisation defect of the DM mutant, it was difficult to conclude whether RNA dimerisation played some role in particle morphogenesis. The aforementioned evidence gathered in HIV-1 tends to suggest that it does not and it seemed more likely that the defect in producing fully mature mutant virions was the consequence of the reduced genome encapsidation. Interestingly, genomic RNA packaging was also found to play a role in particle maturation in SIV_{MAC}, where mutations of SL-1, which led to a decreased packaging and dimerisation efficiency, were reported to affect the production of fully mature particles (Whitney and Wainberg, 2006). Since the HIV-2 SL-1 was also shown to be involved in genome dimerisation and encapsidation (Griffin et al., 2001; McCann and Lever, 1997; present study), it was tempting to draw parallels between HIV-2 and SIV. Indeed, the present report showed that the DM packaging mutant failed to dimerise its genome and to efficiently produce fully infectious mature particles, corroborating the findings in SIV (Whitney and Wainberg, 2006). Furthermore, these results supported the suggestion of a close relationship between the packaging of the HIV-2 genomic RNA - and possibly its dimerisation -, the infectivity and replicative fitness of the virus and the correct assembly and maturation of HIV-2 virions.

6 Concluding remarks

The major focus of this work was to determine the role played by the HIV-2 Psi/SL-1 sequence and structure in genomic RNA dimerisation and to explore the possible connection between genome dimerisation and encapsidation, viral infectivity and particle morphogenesis.

Using a range of Psi/SL-1 mutants, we have determined that a palindrome located in the loop of SL-1 and previously named the HIV-2 DIS was not required for the dimerisation of HIV-2 genomic RNA *in virio*. Instead, a palindromic sequence within the Psi region, *pal*, was shown to be involved. The HIV-2 'DIS' palindrome was not essential for RNA encapsidation and viral replication in T-cells, whereas mutations of *pal* affected both RNA packaging and infectivity.

Further mapping of the sequence and structural elements involved in HIV-2 RNA dimerisation revealed that a purine-rich motif (GGAG) at the 5′ extremity of *pal* was absolutely required for genome dimerisation and viral replication. However, contrary to a previous report proposing that the formation of a stem B structure at the bottom of SL-1 was important for RNA packaging and viral replication, the formation of stem B was not found to affect HIV-2 encapsidation. Viruses bearing mutations that disrupted the formation of stem B showed a reduction in infectivity, which was restored by compensatory mutations. However, this result was only observed in a single-round infectivity assay and the disruption of stem B did not affect the replication kinetics of the virus, suggesting that the absence of stem B may affect early events during infection but that the defect is easily overcome by the virus, possibly due to the redundancy of the

elements involved. The presence of a second purine-rich motif at position 439 of the HIV-2_{ROD} genome did however appear to be important for viral replication as substitution of this GGAG motif for a CCUC resulted in a delayed replication phenotype, and a double mutant of both GGAG sequences showed the most dramatic reduction in viral infectivity. The first motif, located in the Psi region, appears to have a dominant effect as shown by a comparison of the two individual GGAG mutants, SM2 and SM6. The purine-rich motifs present in HIV-2 5′ leader RNA therefore seem to play a role in the replication cycle.

In this study, eleven mutant viruses were analysed alongside wild type HIV-2. A correlation emerged between the capacity to form RNA dimers, the packaging efficiency and the replication fitness of the viruses. Importantly, while all mutants that showed a dimerisation defect also failed to efficiently package their genome, the reverse was not true, suggesting a nonreciprocal relationship between genome dimerisation and encapsidation (see Figure 6.1). These results imply that, in HIV-2, RNA dimer formation is important for encapsidation and that genome dimerisation likely occurs prior to packaging.

Genome dimerisation appears to be directly linked with viral infectivity and the GGAG motif at position 392-395 of the HIV-2_{ROD} genome is critical for both HIV-2 replication and RNA dimer formation, but not for encapsidation. Furthermore, mutants with packaging defects of up to 50 % could establish a productive infection in T-cells, demonstrating that viral replication and encapsidation are not interdependent (Figure 6.1). Besides, viral infectivity was affected independently of the amount of RNA genomes targeted to the cells, suggesting that other aspects of the virus life cycle may be affected either by the mutation introduced or by the reduction in genome dimerisation. The efficacy of first-strand transfer during reverse transcription has been

shown to decrease in HIV-1 mutants bearing mutations in the dimerisation site. Attempts were made to determine whether reverse transcription was affected in the Psi/SL-1 mutants, but these remained unsuccessful.

Long-term culture of SL-1 dimerisation-mutants resulted in the appearance of revertants with compensatory mutations in MA, CA and p6, suggesting a possible implication for either the RNA dimer or the 392-395 GGAG motif in protein synthesis, proteolytic processing, particle assembly and stability.

An effect on the stability of the particle was observed with the DM mutant, which failed to dimerise and package its genome efficiently. Fewer mature particles were observed, together with an increased number of particles containing two cores but the same proportion of immature particles, suggesting that the defect probably occurred after particle release and initiation of particle maturation. An interesting hypothesis is that the interaction of Gag with the RNA genome may prevent the destabilisation of the capsid during the proteolytic processing of the polyproteins. Studies in HIV-1 and SIV tend to suggest that the packaging of RNA rather than the dimeric nature of this RNA is important for particle assembly and stability. This would suggest that the morphogenesis defect observed by electron microscopy for the DM mutant resulted from its packaging defect rather than from its inability to form stable RNA dimers.

Finally, the presence of purine-rich motifs (e.g. GGAG) in the HIV-2 5′ UTR appears to be essential for viral infectivity. Based on *in vitro* studies of HIV-1 Gag-RNA interaction, these motifs represent obvious Gag/NC binding site and it seems likely that the binding of Gag to one or several of these sites would be involved in the regulation of dimerisation, packaging and possibly translation. Unfortunately, our attempts to determine the binding affinity of HIV-2 Gag to the SL-1 mutants were unsuccessful, and this hypothesis could not be verified.

From the results presented in this report, it can be established that the presence of a purine-rich motif at position 392-395 of the HIV-2 genome correlates with RNA dimerisation and viral replication. Nonetheless, the connection with genome encapsidation remains uncertain. While genome dimerisation is required for viral replication and may be important for early steps of the virus life cycle such as reverse transcription, efficient RNA packaging, possibly in combination with the formation of a stable RNA dimer, seems to be essential for particle assembly and morphogenesis (see Figure 6.1). Therefore, it seems likely that these processes may be connected at specific times of the virus life cycle, potentially ensuring the correct regulation of events that is necessary for the production of infectious particles.

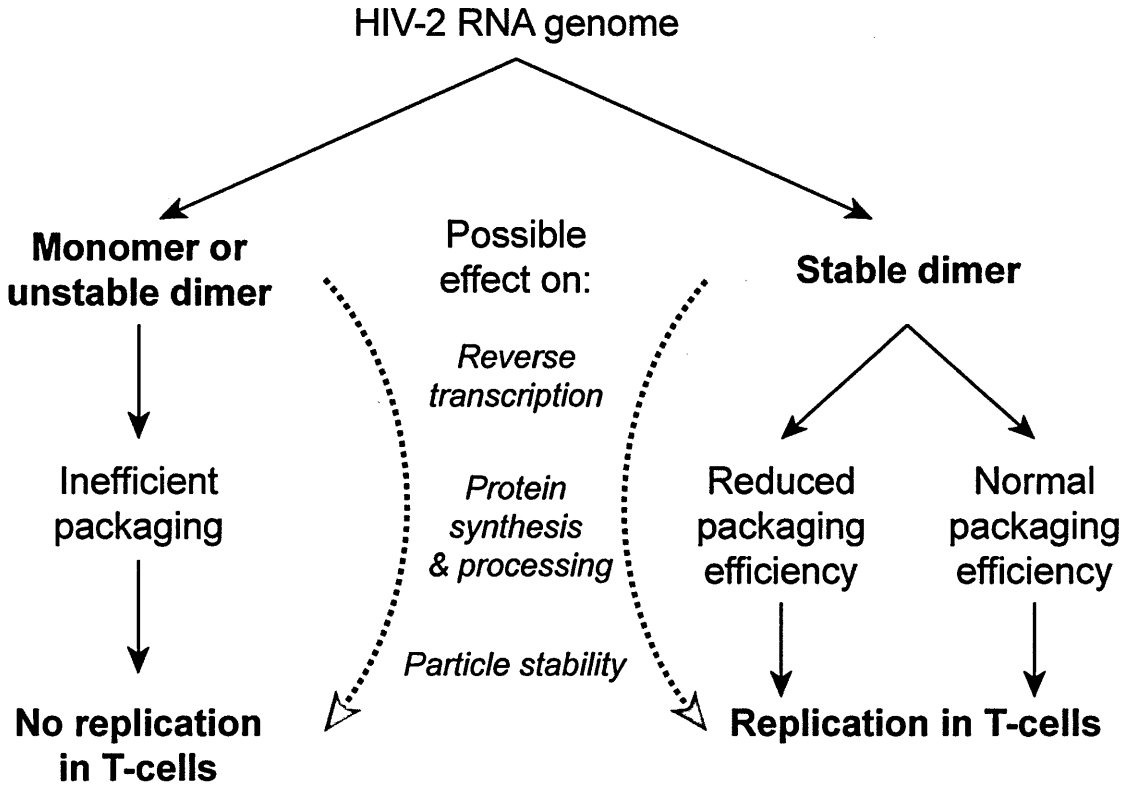


Figure 6.1: Relationship between HIV-2 genome dimerisation, packaging and viral replication.

The association between genome dimerisation, packaging and viral replication is derived from the analysis of a series of Psi/SL-1 mutants, with phenotypes which fulfil the categories described. The dashed lines denote possible sites of action, based on previous findings in other retroviruses, which would be worth investigating in HIV-2.

7 References

Ambion Technical Bulletin # 500 (http://www.ambion.com/techlib/tb/tb_500.html).

HIV Sequence Compendium 2008. (C. Kuiken, T. Leitner, B. Foley, B. Hahn, P. Marx, F. McCutchan, S. Wolinsky, and B. Korber editors. Publisher: Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, New Mexico. LA-UR 08-03719.).

Abada, P., Noble, B., and Cannon, P.M. (2005). Functional domains within the human immunodeficiency virus type 2 envelope protein required to enhance virus production. J Virol *79*, 3627-3638.

Abbink, T.E., and Berkhout, B. (2003). A novel long distance base-pairing interaction in human immunodeficiency virus type 1 RNA occludes the Gag start codon. J Biol Chem *278*, 11601-11611.

Abbink, T.E., and Berkhout, B. (2007). HIV-1 reverse transcription: close encounters between the viral genome and a cellular tRNA. Adv Pharmacol *55*, 99-135.

Abbink, T.E., Ooms, M., Haasnoot, P.C., and Berkhout, B. (2005). The HIV-1 leader RNA conformational switch regulates RNA dimerization but does not regulate mRNA translation. Biochemistry *44*, 9058-9066.

Aiken, C., Konner, J., Landau, N.R., Lenburg, M.E., and Trono, D. (1994). Nef induces CD4 endocytosis: requirement for a critical dileucine motif in the membrane-proximal CD4 cytoplasmic domain. Cell *76*, 853-864.

Aldovini, A., and Young, R.A. (1990). Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. J Virol *64*, 1920-1926.

Alizon, M., Sonigo, P., Barre-Sinoussi, F., Chermann, J.C., Tiollais, P., Montagnier, L., and Wain-Hobson, S. (1984). Molecular cloning of lymphadenopathy-associated virus. Nature *312*, 757-760.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res *25*, 3389-3402.

Amarasinghe, G.K., De Guzman, R.N., Turner, R.B., Chancellor, K.J., Wu, Z.R., and Summers, M.F. (2000a). NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. J Mol Biol *301*, 491-511.

Amarasinghe, G.K., De Guzman, R.N., Turner, R.B., and Summers, M.F. (2000b). NMR structure of stem-loop SL2 of the HIV-1 psi RNA packaging signal reveals a novel A-U-A base-triple platform. J Mol Biol *299*, 145-156.

Amarasinghe, G.K., Zhou, J., Miskimon, M., Chancellor, K.J., McDonald, J.A., Matthews, A.G., Miller, R.R., Rouse, M.D., and Summers, M.F. (2001). Stem-loop SL4 of the HIV-1 Psi RNA Packaging Signal Exhibits Weak Affinity for the Nucleocapsid Protein. Structural Studies and Implications for Genome Recognition. J Mol Biol *314*, 961-970.

Andersen, E.S., Jeeninga, R.E., Damgaard, C.K., Berkhout, B., and Kjems, J. (2003). Dimerization and template switching in the 5' untranslated region between various subtypes of human immunodeficiency virus type 1. J Virol *77*, 3020-3030.

Anderson, E.C. Personal communication.

Anderson, E.C., and Lever, A.M. (2006). Human immunodeficiency virus type 1 Gag polyprotein modulates its own translation. J Virol *80*, 10478-10486.

Anderson, R.M., and May, R.M. (1996). The population biology of the interaction between HIV-1 and HIV-2: coexistence or competitive exclusion? AIDS *10*, 1663-1673.

Arya, S.K., Zamani, M., and Kundra, P. (1998). Human immunodeficiency virus type 2 lentivirus vectors for gene transfer: expression and potential for helper virus-free packaging. Hum Gene Ther *9*, 1371-1380.

Auersperg, N. (1964). Long-Term Cultivation of Hypodiploid Human Tumor Cells. J Natl Cancer Inst *32*, 135-163.

Awang, G., and Sen, D. (1993). Mode of dimerization of HIV-1 genomic RNA. Biochemistry *32*, 11453-11457.

Bacharach, E., and Goff, S.P. (1998). Binding of the human immunodeficiency virus type 1 Gag protein to the viral RNA encapsidation signal in the yeast three-hybrid system. J Virol *72*, 6944-6949.

Baig, T.T., Lanchy, J.M., and Lodmell, J.S. (2007). HIV-2 RNA dimerization is regulated by intramolecular interactions in vitro. RNA *13*, 1341-1354.

Baig, T.T., Lanchy, J.M., and Lodmell, J.S. (2009). Randomization and in vivo selection reveal a GGRG motif essential for packaging human immunodeficiency virus type 2 RNA. J Virol *83*, 802-810.

Balakrishnan, M., Fay, P.J., and Bambara, R.A. (2001). The kissing hairpin sequence promotes recombination within the HIV-I 5' leader region. J Biol Chem *276*, 36482-36492.

Balakrishnan, M., Roques, B.P., Fay, P.J., and Bambara, R.A. (2003). Template dimerization promotes an acceptor invasion-induced transfer mechanism during human immunodeficiency virus type 1 minus-strand synthesis. J Virol *77*, 4710-4721.

Barr, S.D., Smiley, J.R., and Bushman, F.D. (2008). The interferon response inhibits HIV particle production by induction of TRIM22. PLoS Pathog *4*, e1000007.

Barre-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., *et al.* (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science *220*, 868-871.

Basyuk, E., Galli, T., Mougel, M., Blanchard, J.M., Sitbon, M., and Bertrand, E. (2003). Retroviral genomic RNAs are transported to the plasma membrane by endosomal vesicles. Dev Cell *5*, 161-174.

Baudin, F., Marquet, R., Isel, C., Darlix, J.L., Ehresmann, B., and Ehresmann, C. (1993). Functional sites in the 5' region of human immunodeficiency virus type 1 RNA form defined structural domains. J Mol Biol *229*, 382-397.

Beer, B.E., Bailes, E., Sharp, P. and Hirsch, V. (1999). Diversity and Evolution of Primate Lentiviruses. In The Human Retroviruses and AIDS 1999 Compendium, B.F. Carla Kuiken, Beatrice Hahn, Preston Marx, Francince McCutchan, John Mellors, James Mullins, Steven Wolinsky, and Bette Korber, ed. (Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, New Mexico).

Bender, W., Chien, Y.H., Chattopadhyay, S., Vogt, P.K., Gardner, M.B., and Davidson, N. (1978). High-molecular-weight RNAs of AKR, NZB, and wild mouse viruses and avian reticuloendotheliosis virus all have similar dimer structures. J Virol *25*, 888-896.

Bender, W., and Davidson, N. (1976). Mapping of poly(A) sequences in the electron microscope reveals unusual structure of type C oncovirus RNA molecules. Cell *7*, 595-607.

Bennett, E.M., Lever, A.M., and Allen, J.F. (2004). Human immunodeficiency virus type 2 Gag interacts specifically with PRP4, a serine-threonine kinase, and inhibits phosphorylation of splicing factor SF2. J Virol *78*, 11303-11312.

Berkhout, B. (1996). Structure and function of the human immunodeficiency virus leader RNA. Prog Nucleic Acid Res Mol Biol *54*, 1-34.

Berkhout, B., Das, A.T., and van Wamel, J.L. (1998). The native structure of the human immunodeficiency virus type 1 RNA genome is required for the first strand transfer of reverse transcription. Virology *249*, 211-218.

Berkhout, B., Essink, B.B., and Schoneveld, I. (1993). In vitro dimerization of HIV-2 leader RNA in the absence of PuGGAPuA motifs. Faseb J *7*, 181-187.

Berkhout, B., Gatignol, A., Silver, J., and Jeang, K.T. (1990). Efficient trans-activation by the HIV-2 Tat protein requires a duplicated TAR RNA structure. Nucleic Acids Res *18*, 1839-1846.

Berkhout, B., and Jeang, K.T. (1989). trans activation of human immunodeficiency virus type 1 is sequence specific for both the single-stranded bulge and loop of the trans-acting-responsive hairpin: a quantitative analysis. J Virol *63*, 5501-5504.

Berkhout, B., Ooms, M., Beerens, N., Huthoff, H., Southern, E., and Verhoef, K. (2002). In vitro evidence that the untranslated leader of the HIV-1 genome is an RNA checkpoint that regulates multiple functions through conformational changes. J Biol Chem *277*, 19967-19975.

Berkhout, B., and van Wamel, J.L. (1996). Role of the DIS hairpin in replication of human immunodeficiency virus type 1. J Virol *70*, 6723-6732.

Berkhout, B., and van Wamel, J.L. (2000). The leader of the HIV-1 RNA genome forms a compactly folded tertiary structure. RNA *6*, 282-295.

Berkowitz, R.D., and Goff, S.P. (1994). Analysis of binding elements in the human immunodeficiency virus type 1 genomic RNA and nucleocapsid protein. Virology *202*, 233-246.

Berkowitz, R.D., Luban, J., and Goff, S.P. (1993). Specific binding of human immunodeficiency virus type 1 gag polyprotein and nucleocapsid protein to viral RNAs detected by RNA mobility shift assays. J Virol *67*, 7190-7200.

Berkowitz, R.D., Ohagen, A., Hoglund, S., and Goff, S.P. (1995). Retroviral nucleocapsid domains mediate the specific recognition of genomic viral RNAs by chimeric Gag polyproteins during RNA packaging in vivo. J Virol *69*, 6445-6456.

Berry, N., Ariyoshi, K., Jaffar, S., Sabally, S., Corrah, T., Tedder, R., and Whittle, H. (1998). Low peripheral blood viral HIV-2 RNA in individuals with high CD4 percentage differentiates HIV-2 from HIV-1 infection. J Hum Virol *1*, 457-468.

Berry, N., Davis, C., Jenkins, A., Wood, D., Minor, P., Schild, G., Bottiger, M., Holmes, H., and Almond, N. (2001). Vaccine safety. Analysis of oral polio vaccine CHAT stocks. Nature *410*, 1046-1047.

Berry, N., Jaffar, S., Schim van der Loeff, M., Ariyoshi, K., Harding, E., N'Gom, P.T., Dias, F., Wilkins, A., Ricard, D., Aaby, P., *et al.* (2002). Low level viremia and high CD4% predict normal survival in a cohort of HIV type-2-infected villagers. AIDS Res Hum Retroviruses *18*, 1167-1173.

Berry, N., Jenkins, A., Martin, J., Davis, C., Wood, D., Schild, G., Bottiger, M., Holmes, H., Minor, P., and Almond, N. (2005). Mitochondrial DNA and retroviral RNA analyses of archival oral polio vaccine (OPV CHAT) materials: evidence of macaque nuclear sequences confirms substrate identity. Vaccine *23*, 1639-1648.

Betts, M.J., and Russell, R.B. (2003). Amino acid properties and consequences of subsitutions. In Bioinformatics for Geneticists, M.R. Barnes, and I.C. Gray, eds. (Wiley).

Bieniasz, P.D., Grdina, T.A., Bogerd, H.P., and Cullen, B.R. (1999). Recruitment of cyclin T1/P-TEFb to an HIV type 1 long terminal repeat promoter proximal RNA target is both necessary and sufficient for full activation of transcription. Proc Natl Acad Sci USA *96*, 7791-7796.

Bishop, K.N., Verma, M., Kim, E.Y., Wolinsky, S.M., and Malim, M.H. (2008). APOBEC3G inhibits elongation of HIV-1 reverse transcripts. PLoS Pathog *4*, e1000231.

Blaak, H., Brouwer, M., Ran, L.J., de Wolf, F., and Schuitemaker, H. (1998). In vitro replication kinetics of human immunodeficiency virus type 1 (HIV-1) variants in relation to virus load in long-term survivors of HIV-1 infection. J Infect Dis *177*, 600-610.

Blaak, H., van der Ende, M.E., Boers, P.H., Schuitemaker, H., and Osterhaus, A.D. (2006). In vitro replication capacity of HIV-2 variants from long-term aviremic individuals. Virology *353*, 144-154.

Bogerd, H.P., Doehle, B.P., Wiegand, H.L., and Cullen, B.R. (2004). A single amino acid difference in the host APOBEC3G protein controls the primate species specificity of HIV type 1 virion infectivity factor. Proc Natl Acad Sci USA *101*, 3770-3774.

Bogerd, H.P., Echarri, A., Ross, T.M., and Cullen, B.R. (1998). Inhibition of human immunodeficiency virus Rev and human T-cell leukemia virus Rex function, but not Mason-Pfizer monkey virus constitutive transport element activity, by a mutant human nucleoporin targeted to Crm1. J Virol *72*, 8627-8635.

Bonnet-Mathoniere, B., Girard, P.M., Muriaux, D., and Paoletti, J. (1996). Nucleocapsid protein 10 activates dimerization of the RNA of Moloney murine leukaemia virus in vitro. Eur J Biochem *238*, 129-135.

Bouamr, F., Scarlata, S., and Carter, C. (2003). Role of myristylation in HIV-1 Gag assembly. Biochemistry *42*, 6408-6417.

Bour, S., Akari, H., Miyagi, E., and Strebel, K. (2003). Naturally occurring amino acid substitutions in the HIV-2 ROD envelope glycoprotein regulate its ability to augment viral particle release. Virology *309*, 85-98.

Bour, S., Schubert, U., Peden, K., and Strebel, K. (1996). The envelope glycoprotein of human immunodeficiency virus type 2 enhances viral particle release: a Vpu-like factor? J Virol *70*, 820-829.

Bour, S., and Strebel, K. (1996). The human immunodeficiency virus (HIV) type 2 envelope protein is a functional complement to HIV type 1 Vpu that enhances particle release of heterologous retroviruses. J Virol *70*, 8285-8300.

Braaten, D., Franke, E.K., and Luban, J. (1996). Cyclophilin A is required for an early step in the life cycle of human immunodeficiency virus type 1 before the initiation of reverse transcription. J Virol *70*, 3551-3560.

Brasey, A., Lopez-Lastra, M., Ohlmann, T., Beerens, N., Berkhout, B., Darlix, J.L., and Sonenberg, N. (2003). The leader of human immunodeficiency virus type 1 genomic RNA harbors an internal ribosome entry segment that is active during the G2/M phase of the cell cycle. J Virol *77*, 3939-3949.

Brierley, I., and Dos Ramos, F.J. (2006). Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. Virus Res *119*, 29-42.

Briggs, J.A. (2008). Structure and arrangement of Gag in immature retroviruses. Abstract presented at "The retrovirus assembly meeting", October 4-8 2008. (Prague, Czech Republic).

Briggs, J.A., Wilk, T., Welker, R., Krausslich, H.G., and Fuller, S.D. (2003). Structural organization of authentic, mature HIV-1 virions and cores. EMBO J *22*, 1707-1715.

Brun-Vezinet, F., Rey, M.A., Katlama, C., Girard, P.M., Roulot, D., Yeni, P., Lenoble, L., Clavel, F., Alizon, M., Gadelle, S., *et al.* (1987). Lymphadenopathy-associated virus

type 2 in AIDS and AIDS-related complex. Clinical and virological features in four patients. Lancet *1*, 128-132.

Bryant, M., and Ratner, L. (1990). Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. Proc Natl Acad Sci USA *87*, 523-527.

Buck, C.B., Shen, X., Egan, M.A., Pierson, T.C., Walker, C.M., and Siliciano, R.F. (2001). The human immunodeficiency virus type 1 gag gene encodes an internal ribosome entry site. J Virol *75*, 181-191.

Bukrinsky, M.I., Haggerty, S., Dempsey, M.P., Sharova, N., Adzhubel, A., Spitz, L., Lewis, P., Goldfarb, D., Emerman, M., and Stevenson, M. (1993a). A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells. Nature *365*, 666-669.

Bukrinsky, M.I., Sharova, N., McDonald, T.L., Pushkarskaya, T., Tarpley, W.G., and Stevenson, M. (1993b). Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. Proc Natl Acad Sci USA *90*, 6125-6129.

Burniston, M.T., Cimarelli, A., Colgan, J., Curtis, S.P., and Luban, J. (1999). Human immunodeficiency virus type 1 Gag polyprotein multimerization requires the nucleocapsid domain and RNA and is promoted by the capsid- dimer interface and the basic region of matrix protein. J Virol *73*, 8527-8540.

Burton, G.F., Keele, B.F., Estes, J.D., Thacker, T.C., and Gartner, S. (2002). Follicular dendritic cell contributions to HIV pathogenesis. Semin Immunol *14*, 275-284.

Busschots, K., Vercammen, J., Emiliani, S., Benarous, R., Engelborghs, Y., Christ, F., and Debyser, Z. (2005). The interaction of LEDGF/p75 with integrase is lentivirus-specific and promotes DNA binding. J Biol Chem *280*, 17841-17847.

Campbell, S., and Rein, A. (1999). In vitro assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. J Virol *73*, 2270-2279.

Campbell, S., and Vogt, V.M. (1995). Self-assembly in vitro of purified CA-NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1. J Virol *69*, 6487-6497.

Campbell, S., and Vogt, V.M. (1997). In vitro assembly of virus-like particles with Rous sarcoma virus Gag deletion mutants: identification of the p10 domain as a morphological determinant in the formation of spherical particles. J Virol *71*, 4425-4435.

Cannon, P. Personal communication.

Cao, Y., Qin, L., Zhang, L., Safrit, J., and Ho, D.D. (1995). Virologic and immunologic characterization of long-term survivors of human immunodeficiency virus type 1 infection. N Engl J Med *332*, 201-208.

CDC (1996). Identification of HIV-1 group O infection--Los Angeles county, California, 1996. In MMWR Morb Mortal Wkly Rep, pp. 561-565.

Chang, Y.F., Wang, S.M., Huang, K.J., and Wang, C.T. (2007). Mutations in capsid major homology region affect assembly and membrane affinity of HIV-1 Gag. J Mol Biol *370*, 585-597.

Chen, Z., Luckay, A., Sodora, D.L., Telfer, P., Reed, P., Gettie, A., Kanu, J.M., Sadek, R.F., Yee, J., Ho, D.D., *et al.* (1997). Human immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range of simian immunodeficiency virus-infected sooty mangabeys. J Virol *71*, 3953-3960.

Chen, Z., Telfier, P., Gettie, A., Reed, P., Zhang, L., Ho, D.D., and Marx, P.A. (1996). Genetic characterization of new West African simian immunodeficiency virus SIVsm: geographic clustering of household-derived SIV strains with human immunodeficiency virus type 2 subtypes and genetically diverse viruses from a single feral sooty mangabey troop. J Virol *70*, 3617-3627.

Chin, M.P., Lee, S.K., Chen, J., Nikolaitchik, O.A., Powell, D.A., Fivash, M.J., Jr., and Hu, W.S. (2008). Long-range recombination gradient between HIV-1 subtypes B and C variants caused by sequence differences in the dimerization initiation signal region. J Mol Biol *377*, 1324-1333.

Cimarelli, A., Sandin, S., Hoglund, S., and Luban, J. (2000). Basic residues in human immunodeficiency virus type 1 nucleocapsid promote virion assembly via interaction with RNA. J Virol *74*, 3046-3057.

Clavel, F., Guetard, D., Brun-Vezinet, F., Chamaret, S., Rey, M.A., Santos-Ferreira, M.O., Laurent, A.G., Dauguet, C., Katlama, C., Rouzioux, C., *et al.* (1986a). Isolation of a new human retrovirus from West African patients with AIDS. Science *233*, 343-346.

Clavel, F., Guyader, M., Guetard, D., Salle, M., Montagnier, L., and Alizon, M. (1986b). Molecular cloning and polymorphism of the human immune deficiency virus type 2. Nature *324*, 691-695.

Clavel, F., and Orenstein, J.M. (1990). A mutant of human immunodeficiency virus with reduced RNA packaging and abnormal particle morphology. J Virol *64*, 5230-5234.

Clever, J., Sassetti, C., and Parslow, T.G. (1995). RNA secondary structure and binding sites for gag gene products in the 5' packaging signal of human immunodeficiency virus type 1. J Virol *69*, 2101-2109.

Clever, J.L., and Parslow, T.G. (1997). Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. J Virol *71*, 3407-3414.

Clever, J.L., Wong, M.L., and Parslow, T.G. (1996). Requirements for kissing-loop-mediated dimerization of human immunodeficiency virus RNA. J Virol *70*, 5902-5908.

Coffin, J., Haase, A., Levy, J.A., Montagnier, L., Oroszlan, S., Teich, N., Temin, H., Toyoshima, K., Varmus, H., Vogt, P., *et al.* (1986). What to call the AIDS virus? Nature *321*, 10.

Coffin, J.M., Hughes, S.H., and Varmus, H.E., eds. (1997). Retroviruses, 1 edn (Cold Spring Harbor Laboratory Press).

Collins, K.L., Chen, B.K., Kalams, S.A., Walker, B.D., and Baltimore, D. (1998). HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. Nature *391*, 397-401.

Costa, P., Rusconi, S., Fogli, M., Mavilio, D., Murdaca, G., Puppo, F., Mingari, M.C., Galli, M., Moretta, L., and De Maria, A. (2003). Low expression of inhibitory natural killer receptors in CD8 cytotoxic T lymphocytes in long-term non-progressor HIV-1-infected patients. AIDS *17*, 257-260.

Covey, S.N. (1986). Amino acid sequence homology in gag region of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus. Nucleic Acids Res *14*, 623-633.

Cruceanu, M., Urbaneja, M.A., Hixson, C.V., Johnson, D.G., Datta, S.A., Fivash, M.J., Stephen, A.G., Fisher, R.J., Gorelick, R.J., Casas-Finet, J.R., *et al.* (2006). Nucleic acid binding and chaperone properties of HIV-1 Gag and nucleocapsid proteins. Nucleic Acids Res *34*, 593-605.

Cucchiarini, M., Cagnon, L., Giordanengo, V., Doglio, A., and Lefebvre, J.C. (1995). Induction by human immunodeficiency viruses types 1 and 2 of degradation of CD4 but not of a CD4 mutant unable to bind viral envelope glycoproteins. J Acquir Immune Defic Syndr Hum Retrovirol *8*, 427-436.

Cullen, B.R. (1986). Trans-activation of human immunodeficiency virus occurs via a bimodal mechanism. Cell *46*, 973-982.

Cullen, B.R. (1998). HIV-1 auxiliary proteins: making connections in a dying cell. Cell *93*, 685-692.

D'Souza, V., and Summers, M.F. (2004). Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus. Nature *431*, 586-590.

D'Souza, V., and Summers, M.F. (2005). How retroviruses select their genomes. Nat Rev Microbiol *3*, 643-655.

Dalgleish, A.G., Beverley, P.C., Clapham, P.R., Crawford, D.H., Greaves, M.F., and Weiss, R.A. (1984). The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. Nature *312*, 763-767.

Damgaard, C.K., Andersen, E.S., Knudsen, B., Gorodkin, J., and Kjems, J. (2004). RNA interactions in the 5' region of the HIV-1 genome. J Mol Biol *336*, 369-379.

Damgaard, C.K., Dyhr-Mikkelsen, H., and Kjems, J. (1998). Mapping the RNA binding sites for human immunodeficiency virus type-1 gag and NC proteins within the complete HIV-1 and -2 untranslated leader regions. Nucleic Acids Res *26*, 3667-3676.

Damond, F., Gueudin, M., Pueyo, S., Farfara, I., Robertson, D.L., Descamps, D., Chene, G., Matheron, S., Campa, P., Brun-Vezinet, F., *et al.* (2002). Plasma RNA viral load in human immunodeficiency virus type 2 subtype A and subtype B infections. J Clin Microbiol *40*, 3654-3659.

Damond, F., Worobey, M., Campa, P., Farfara, I., Colin, G., Matheron, S., Brun-Vezinet, F., Robertson, D.L., and Simon, F. (2004). Identification of a highly divergent HIV type 2 and proposal for a change in HIV type 2 classification. AIDS Res Hum Retroviruses *20*, 666-672.

Daniel, M.D., Letvin, N.L., King, N.W., Kannagi, M., Sehgal, P.K., Hunt, R.D., Kanki, P.J., Essex, M., and Desrosiers, R.C. (1985). Isolation of T-cell tropic HTLV-III-like retrovirus from macaques. Science *228*, 1201-1204.

Dannull, J., Surovoy, A., Jung, G., and Moelling, K. (1994). Specific binding of HIV-1 nucleocapsid protein to PSI RNA in vitro requires N-terminal zinc finger and flanking basic amino acid residues. EMBO J *13*, 1525-1533.

Darlix, J.L., Gabus, C., Nugeyre, M.T., Clavel, F., and Barre-Sinoussi, F. (1990). Cis elements and trans-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. J Mol Biol *216*, 689-699.

Dayton, A.I., Sodroski, J.G., Rosen, C.A., Goh, W.C., and Haseltine, W.A. (1986). The trans-activator gene of the human T cell lymphotropic virus type III is required for replication. Cell *44*, 941-947.

De Cock, K.M., Adjorlolo, G., Ekpini, E., Sibailly, T., Kouadio, J., Maran, M., Brattegaard, K., Vetter, K.M., Doorly, R., and Gayle, H.D. (1993). Epidemiology and transmission of HIV-2. Why there is no HIV-2 pandemic. JAMA *270*, 2083-2086.

De Guzman, R.N., Wu, Z.R., Stalling, C.C., Pappalardo, L., Borer, P.N., and Summers, M.F. (1998). Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. Science *279*, 384-388.

De Silva, T.I., Cotten, M., and Rowland-Jones, S.L. (2008). HIV-2: the forgotten AIDS virus. Trends Microbiol *16*, 588-595.

Deacon, N.J., Tsykin, A., Solomon, A., Smith, K., Ludford-Menting, M., Hooker, D.J., McPhee, D.A., Greenway, A.L., Ellett, A., Chatfield, C., *et al.* (1995). Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients. Science *270*, 988-991.

Dehart, J.L., and Planelles, V. (2008). Human immunodeficiency virus type 1 Vpr links proteasomal degradation and checkpoint activation. J Virol *82*, 1066-1072.

Delelis, O., Carayon, K., Saib, A., Deprez, E., and Mouscadet, J.F. (2008). Integrase and integration: biochemical activities of HIV-1 integrase. Retrovirology *5*, 114.

Demarchi, F., D'Agaro, P., Falaschi, A., and Giacca, M. (1992). Probing protein-DNA interactions at the long terminal repeat of human immunodeficiency virus type 1 by in vivo footprinting. J Virol *66*, 2514-2518.

Deneka, M., Pelchen-Matthews, A., Byland, R., Ruiz-Mateos, E., and Marsh, M. (2007). In macrophages, HIV-1 assembles into an intracellular plasma membrane domain containing the tetraspanins CD81, CD9, and CD53. J Cell Biol *177*, 329-341.

Deng, H., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhart, M., Di Marzio, P., Marmon, S., Sutton, R.E., Hill, C.M., *et al.* (1996). Identification of a major co-receptor for primary isolates of HIV-1. Nature *381*, 661-666.

Dirac, A.M., Huthoff, H., Kjems, J., and Berkhout, B. (2001). The dimer initiation site hairpin mediates dimerization of the human immunodeficiency virus, type 2 RNA genome. J Biol Chem *276*, 32345-32352.

Dirac, A.M., Huthoff, H., Kjems, J., and Berkhout, B. (2002). Regulated HIV-2 RNA dimerization by means of alternative RNA conformations. Nucleic Acids Res *30*, 2647-2655.

Dong, X., Li, H., Derdowski, A., Ding, L., Burnett, A., Chen, X., Peters, T.R., Dermody, T.S., Woodruff, E., Wang, J.J., *et al.* (2005). AP-3 directs the intracellular trafficking of HIV-1 Gag and plays a key role in particle assembly. Cell *120*, 663-674.

Dorfman, T., Bukovsky, A., Ohagen, A., Hoglund, S., and Gottlinger, H.G. (1994a). Functional domains of the capsid protein of human immunodeficiency virus type 1. J Virol *68*, 8180-8187.

Dorfman, T., Luban, J., Goff, S.P., Haseltine, W.A., and Gottlinger, H.G. (1993). Mapping of functionally important residues of a cysteine-histidine box in the human immunodeficiency virus type 1 nucleocapsid protein. J Virol *67*, 6159-6169.

Dorfman, T., Mammano, F., Haseltine, W.A., and Gottlinger, H.G. (1994b). Role of the matrix protein in the virion association of the human immunodeficiency virus type 1 envelope glycoprotein. J Virol *68*, 1689-1696.

Dorman, N., and Lever, A. (2000). Comparison of viral genomic RNA sorting mechanisms in human immunodeficiency virus type 1 (HIV-1), HIV-2, and Moloney murine leukemia virus. J Virol *74*, 11413-11417.

Dorman, N.M. (2000). HIV-1 packaging dynamics and inhibition by transcripts containing packaging signal sequences. Ph.D. Department of Medicine: University of Cambridge.

Dragic, T., Litwin, V., Allaway, G.P., Martin, S.R., Huang, Y., Nagashima, K.A., Cayanan, C., Maddon, P.J., Koup, R.A., Moore, J.P., *et al.* (1996). HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. Nature *381*, 667-673.

Duvall, M.G., Jaye, A., Dong, T., Brenchley, J.M., Alabi, A.S., Jeffries, D.J., van der Sande, M., Togun, T.O., McConkey, S.J., Douek, D.C., *et al.* (2006). Maintenance of HIV-specific CD4+ T cell help distinguishes HIV-2 from HIV-1 infection. J Immunol *176*, 6973-6981.

Duvall, M.G., Lore, K., Blaak, H., Ambrozak, D.A., Adams, W.C., Santos, K., Geldmacher, C., Mascola, J.R., McMichael, A.J., Jaye, A., *et al.* (2007). Dendritic cells are less susceptible to human immunodeficiency virus type 2 (HIV-2) infection than to HIV-1 infection. J Virol *81*, 13486-13498.

Duvall, M.G., Precopio, M.L., Ambrozak, D.A., Jaye, A., McMichael, A.J., Whittle, H.C., Roederer, M., Rowland-Jones, S.L., and Koup, R.A. (2008). Polyfunctional T cell responses are a hallmark of HIV-2 infection. Eur J Immunol *38*, 350-363.

Embretson, J., Zupancic, M., Ribas, J.L., Burke, A., Racz, P., Tenner-Racz, K., and Haase, A.T. (1993). Massive covert infection of helper T lymphocytes and macrophages by HIV during the incubation period of AIDS. Nature *362*, 359-362.

Endres, M.J., Clapham, P.R., Marsh, M., Ahuja, M., Turner, J.D., McKnight, A., Thomas, J.F., Stoebenau-Haggarty, B., Choe, S., Vance, P.J., *et al.* (1996). CD4-independent infection by HIV-2 is mediated by fusin/CXCR4. Cell *87*, 745-756.

Engelman, A., and Cherepanov, P. (2008). The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication. PLoS Pathog *4*, e1000046.

Ennifar, E., Walter, P., Ehresmann, B., Ehresmann, C., and Dumas, P. (2001). Crystal structures of coaxially stacked kissing complexes of the HIV-1 RNA dimerization initiation site. Nat Struct Biol *8*, 1064-1068.

Ennifar, E., Yusupov, M., Walter, P., Marquet, R., Ehresmann, B., Ehresmann, C., and Dumas, P. (1999). The crystal structure of the dimerization initiation site of genomic HIV-1 RNA reveals an extended duplex with two adenine bulges. Structure Fold Des *7*, 1439-1449.

Erickson-Viitanen, S., Manfredi, J., Viitanen, P., Tribe, D.E., Tritch, R., Hutchison, C.A., 3rd, Loeb, D.D., and Swanstrom, R. (1989). Cleavage of HIV-1 gag polyprotein synthesized in vitro: sequential cleavage by the viral protease. AIDS Res Hum Retroviruses *5*, 577-591.

Facke, M., Janetzko, A., Shoeman, R.L., and Krausslich, H.G. (1993). A large deletion in the matrix domain of the human immunodeficiency virus gag gene redirects virus particle assembly from the plasma membrane to the endoplasmic reticulum. J Virol *67*, 4972-4980.

Fackler, O.T., Alcover, A., and Schwartz, O. (2007). Modulation of the immunological synapse: a key to HIV-1 pathogenesis? Nat Rev Immunol *7*, 310-317.

Fauquet, C., Mayo, M.A., Maniloff, J., Desselberger, U., and Ball, L.A., eds. (2005). Virus taxonomy VIIIth report of the International Committee on Taxonomy of Viruses (Amsterdam, Elsevier Academic Press).

Feng, S., and Holland, E.C. (1988). HIV-1 tat trans-activation requires the loop sequence within tar. Nature *334*, 165-167.

Feng, Y., Broder, C.C., Kennedy, P.E., and Berger, E.A. (1996a). HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. Science *272*, 872-877.

Feng, Y.X., Campbell, S., Harvin, D., Ehresmann, B., Ehresmann, C., and Rein, A. (1999). The human immunodeficiency virus type 1 Gag polyprotein has nucleic acid chaperone activity: possible role in dimerization of genomic RNA and placement of tRNA on the primer binding site. J Virol *73*, 4251-4256.

Feng, Y.X., Copeland, T.D., Henderson, L.E., Gorelick, R.J., Bosche, W.J., Levin, J.G., and Rein, A. (1996b). HIV-1 nucleocapsid protein induces "maturation" of dimeric retroviral RNA in vitro. Proc Natl Acad Sci USA *93*, 7577-7581.

Fenrick, R., Malim, M.H., Hauber, J., Le, S.Y., Maizel, J., and Cullen, B.R. (1989). Functional analysis of the Tat trans activator of human immunodeficiency virus type 2. J Virol *63*, 5006-5012.

Finzi, A., Orthwein, A., Mercier, J., and Cohen, E.A. (2007). Productive human immunodeficiency virus type 1 assembly takes place at the plasma membrane. J Virol *81*, 7476-7490.

Fisher, R.D., Chung, H.Y., Zhai, Q., Robinson, H., Sundquist, W.I., and Hill, C.P. (2007). Structural and biochemical studies of ALIX/AIP1 and its role in retrovirus budding. Cell *128*, 841-852.

Fletcher, T.M., 3rd, Brichacek, B., Sharova, N., Newman, M.A., Stivahtis, G., Sharp, P.M., Emerman, M., Hahn, B.H., and Stevenson, M. (1996). Nuclear import and cell cycle arrest functions of the HIV-1 Vpr protein are encoded by two separate genes in HIV-2/SIV(SM). EMBO J *15*, 6155-6165.

Fouchier, R.A., and Malim, M.H. (1999). Nuclear import of human immunodeficiency virus type-1 preintegration complexes. Adv Virus Res *52*, 275-299.

Fouchier, R.A., Meyer, B.E., Simon, J.H., Fischer, U., and Malim, M.H. (1997). HIV-1 infection of non-dividing cells: evidence that the amino-terminal basic region of the viral matrix protein is important for Gag processing but not for post-entry nuclear import. EMBO J *16*, 4531-4539.

Franchini, G., Gurgo, C., Guo, H.G., Gallo, R.C., Collalti, E., Fargnoli, K.A., Hall, L.F., Wong-Staal, F., and Reitz, M.S., Jr. (1987). Sequence of simian immunodeficiency virus and its relationship to the human immunodeficiency viruses. Nature *328*, 539-543.

Franchini, G., Rusche, J.R., O'Keeffe, T.J., and Wong-Staal, F. (1988). The human immunodeficiency virus type 2 (HIV-2) contains a novel gene encoding a 16 kD protein associated with mature virions. AIDS Res Hum Retroviruses *4*, 243-250.

Franke, E.K., and Luban, J. (1996). Inhibition of HIV-1 replication by cyclosporine A or related compounds correlates with the ability to disrupt the Gag-cyclophilin A interaction. Virology *222*, 279-282.

Franke, E.K., Yuan, H.E., Bossolt, K.L., Goff, S.P., and Luban, J. (1994a). Specificity and sequence requirements for interactions between various retroviral Gag proteins. J Virol *68*, 5300-5305.

Franke, E.K., Yuan, H.E., and Luban, J. (1994b). Specific incorporation of cyclophilin A into HIV-1 virions. Nature *372*, 359-362.

Freed, E.O. (1998). HIV-1 gag proteins: diverse functions in the virus life cycle. Virology *251*, 1-15.

Freed, E.O., and Martin, M.A. (1995). Virion incorporation of envelope glycoproteins with long but not short cytoplasmic tails is blocked by specific, single amino acid substitutions in the human immunodeficiency virus type 1 matrix. J Virol *69*, 1984-1989.

Freed, E.O., and Martin, M.A. (1996). Domains of the human immunodeficiency virus type 1 matrix and gp41 cytoplasmic tail required for envelope incorporation into virions. J Virol *70*, 341-351.

Freed, E.O., Myers, D.J., and Risser, R. (1989). Mutational analysis of the cleavage sequence of the human immunodeficiency virus type 1 envelope glycoprotein precursor gp160. J Virol *63*, 4670-4675.

Friedman-Kien, A.E., Laubenstein, L.J., Rubinstein, P., Buimovici-Klein, E., Marmor, M., Stahl, R., Spigland, I., Kim, K.S., and Zolla-Pazner, S. (1982). Disseminated Kaposi's sarcoma in homosexual men. Ann Intern Med *96*, 693-700.

Fu, W., Dang, Q., Nagashima, K., Freed, E.O., Pathak, V.K., and Hu, W.S. (2006). Effects of Gag mutation and processing on retroviral dimeric RNA maturation. J Virol *80*, 1242-1249.

Fu, W., Gorelick, R.J., and Rein, A. (1994). Characterization of human immunodeficiency virus type 1 dimeric RNA from wild-type and protease-defective virions. J Virol *68*, 5013-5018.

Fu, W., Prasad, V.V., Chen, J., Nikolaitchik, O., and Hu, W.S. (2007). Molecular mechanisms of simian immunodeficiency virus SIV(agm) RNA encapsidation. Virology *363*, 210-219.

Fu, W., and Rein, A. (1993). Maturation of dimeric viral RNA of Moloney murine leukemia virus. J Virol *67*, 5443 - 5449.

Fujii, K., Hurley, J.H., and Freed, E.O. (2007). Beyond Tsg101: the role of Alix in 'ESCRTing' HIV-1. Nat Rev Microbiol *5*, 912-916.

Fuller, S.D., Wilk, T., Gowen, B.E., Krausslich, H.G., and Vogt, V.M. (1997). Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. Curr Biol *7*, 729-738.

Gaines, H., von Sydow, M.A., von Stedingk, L.V., Biberfeld, G., Bottiger, B., Hansson, L.O., Lundbergh, P., Sonnerborg, A.B., Wasserman, J., and Strannegaard, O.O. (1990). Immunological changes in primary HIV-1 infection. AIDS *4*, 995-999.

Gallay, P., Swingler, S., Aiken, C., and Trono, D. (1995a). HIV-1 infection of nondividing cells: C-terminal tyrosine phosphorylation of the viral matrix protein is a key regulator. Cell *80*, 379-388.

Gallay, P., Swingler, S., Song, J., Bushman, F., and Trono, D. (1995b). HIV nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase. Cell *83*, 569-576.

Gallo, R.C., Salahuddin, S.Z., Popovic, M., Shearer, G.M., Kaplan, M., Haynes, B.F., Palker, T.J., Redfield, R., Oleske, J., Safai, B., *et al.* (1984). Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science *224*, 500-503.

Gallo, R.C., Sarin, P.S., Gelmann, E.P., Robert-Guroff, M., Richardson, E., Kalyanaraman, V.S., Mann, D., Sidhu, G.D., Stahl, R.E., Zolla-Pazner, S., *et al.* (1983). Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). Science *220*, 865-867.

Gamble, T.R., Vajdos, F.F., Yoo, S., Worthylake, D.K., Houseweart, M., Sundquist, W.I., and Hill, C.P. (1996). Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. Cell *87*, 1285-1294.

Gamble, T.R., Yoo, S., Vajdos, F.F., von Schwedler, U.K., Worthylake, D.K., Wang, H., McCutcheon, J.P., Sundquist, W.I., and Hill, C.P. (1997). Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. Science *278*, 849-853.

Ganser-Pornillos, B.K., von Schwedler, U.K., Stray, K.M., Aiken, C., and Sundquist, W.I. (2004). Assembly properties of the human immunodeficiency virus type 1 CA protein. J Virol *78*, 2545-2552.

Ganser-Pornillos, B.K., Yeager, M., and Sundquist, W.I. (2008). The structural biology of HIV assembly. Curr Opin Struct Biol *18*, 203-217.

Ganser, B.K., Li, S., Klishko, V.Y., Finch, J.T., and Sundquist, W.I. (1999). Assembly and analysis of conical models for the HIV-1 core. Science *283*, 80-83.

Gao, F., Bailes, E., Robertson, D.L., Chen, Y., Rodenburg, C.M., Michael, S.F., Cummins, L.B., Arthur, L.O., Peeters, M., Shaw, G.M., *et al.* (1999). Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature *397*, 436-441.

Gao, F., Yue, L., Robertson, D.L., Hill, S.C., Hui, H., Biggar, R.J., Neequaye, A.E., Whelan, T.M., Ho, D.D., Shaw, G.M., *et al.* (1994). Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. J Virol *68*, 7433-7447.

Gao, F., Yue, L., White, A.T., Pappas, P.G., Barchue, J., Hanson, A.P., Greene, B.M., Sharp, P.M., Shaw, G.M., and Hahn, B.H. (1992). Human infection by genetically diverse SIVsm-related HIV-2 in west Africa. Nature *358*, 495-499.

Garbitt, R.A., Albert, J.A., Kessler, M.D., and Parent, L.J. (2001). Trans-acting inhibition of genomic RNA dimerization by Rous sarcoma virus matrix mutants. J Virol *75*, 260-268.

Garrus, J.E., von Schwedler, U.K., Pornillos, O.W., Morham, S.G., Zavitz, K.H., Wang, H.E., Wettstein, D.A., Stray, K.M., Cote, M., Rich, R.L., *et al.* (2001). Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. Cell *107*, 55-65.

Garzino-Demo, A., Gallo, R.C., and Arya, S.K. (1995). Human immunodeficiency virus type 2 (HIV-2): packaging signal and associated negative regulatory element. Hum Gene Ther *6*, 177-184.

Gelderblom, H.R. (1991). Assembly and morphology of HIV: potential effect of structure on viral function. AIDS *5*, 617-637.

Ghys, P.D., Fransen, K., Diallo, M.O., Ettiegne-Traore, V., Coulibaly, I.M., Yeboue, K.M., Kalish, M.L., Maurice, C., Whitaker, J.P., Greenberg, A.E., *et al.* (1997). The associations between cervicovaginal HIV shedding, sexually transmitted diseases and immunosuppression in female sex workers in Abidjan, Cote d'Ivoire. AIDS *11*, F85-93.

Girard, F., Barbault, F., Gouyette, C., Huynh-Dinh, T., Paoletti, J., and Lancelot, G. (1999). Dimer initiation sequence of HIV-1Lai genomic RNA: NMR solution structure of the extended duplex. J Biomol Struct Dyn *16*, 1145-1157.

Gluzman, Y. (1981). SV40-transformed simian cells support the replication of early SV40 mutants. Cell *23*, 175-182.

Goff, A., Ehrlich, L.S., Cohen, S.N., and Carter, C.A. (2003). Tsg101 control of human immunodeficiency virus type 1 Gag trafficking and release. J Virol *77*, 9173-9182.

Gomez, C.Y., and Hope, T.J. (2006). Mobility of human immunodeficiency virus type 1 Pr55Gag in living cells. J Virol *80*, 8796-8806.

Gordon, S.N., Dunham, R.M., Engram, J.C., Estes, J., Wang, Z., Klatt, N.R., Paiardini, M., Pandrea, I.V., Apetrei, C., Sodora, D.L., *et al.* (2008). Short-lived infected cells support virus replication in sooty mangabeys naturally infected with simian immunodeficiency virus: implications for AIDS pathogenesis. J Virol *82*, 3725-3735.

Gorelick, R.J., Chabot, D.J., Rein, A., Henderson, L.E., and Arthur, L.O. (1993). The two zinc fingers in the human immunodeficiency virus type 1 nucleocapsid protein are not functionally equivalent. J Virol *67*, 4027-4036.

Gorelick, R.J., Henderson, L.E., Hanser, J.P., and Rein, A. (1988). Point mutants of Moloney murine leukemia virus that fail to package viral RNA: evidence for specific RNA recognition by a "zinc finger-like" protein sequence. Proc Natl Acad Sci USA *85*, 8420-8424.

Gorelick, R.J., Nigida, S.M., Jr., Bess, J.W., Jr., Arthur, L.O., Henderson, L.E., and Rein, A. (1990). Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. J Virol *64*, 3207-3211.

Gottlieb, G.S., Sow, P.S., Hawes, S.E., Ndoye, I., Redman, M., Coll-Seck, A.M., Faye-Niang, M.A., Diop, A., Kuypers, J.M., Critchlow, C.W., *et al.* (2002). Equal plasma viral loads predict a similar rate of CD4+ T cell decline in human immunodeficiency virus (HIV) type 1- and HIV-2-infected individuals from Senegal, West Africa. J Infect Dis *185*, 905-914.

Gottlieb, M.S., Schroff, R., Schanker, H.M., Weisman, J.D., Fan, P.T., Wolf, R.A., and Saxon, A. (1981). Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. N Engl J Med *305*, 1425-1431.

Gottlinger, H.G. (2001). The HIV-1 assembly machine. AIDS *15 Suppl 5*, S13-20.

Gottlinger, H.G., Dorfman, T., Sodroski, J.G., and Haseltine, W.A. (1991). Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. Proc Natl Acad Sci USA *88*, 3195-3199.

Goujon, C., Riviere, L., Jarrosson-Wuilleme, L., Bernaud, J., Rigal, D., Darlix, J.L., and Cimarelli, A. (2007). SIVSM/HIV-2 Vpx proteins promote retroviral escape from a proteasome-dependent restriction pathway present in human dendritic cells. Retrovirology *4*, 2.

Gousset, K., Ablan, S.D., Coren, L.V., Ono, A., Soheilian, F., Nagashima, K., Ott, D.E., and Freed, E.O. (2008). Real-time visualization of HIV-1 GAG trafficking in infected macrophages. PLoS Pathog *4*, e1000015.

Granelli-Piperno, A., Finkel, V., Delgado, E., and Steinman, R.M. (1999). Virus replication begins in dendritic cells during the transmission of HIV-1 from mature dendritic cells to T cells. Curr Biol *9*, 21-29.

Greatorex, J. (2004). The retroviral RNA dimer linkage: different structures may reflect different roles. Retrovirology *1*, 22.

Greatorex, J., and Lever, A. (1998). Retroviral RNA dimer linkage. J Gen Virol *79*, 2877-2882.

Green, L.M., and Berg, J.M. (1989). A retroviral Cys-Xaa2-Cys-Xaa4-His-Xaa4-Cys peptide binds metal ions: spectroscopic studies and a proposed three-dimensional structure. Proc Natl Acad Sci USA *86*, 4047-4051.

Grief, C., Hockley, D.J., Fromholc, C.E., and Kitchin, P.A. (1989). The morphology of simian immunodeficiency virus as shown by negative staining electron microscopy. J Gen Virol *70 (Pt 8)*, 2215-2219.

Griffin, S.D., Allen, J.F., and Lever, A.M. (2001). The major human immunodeficiency virus type 2 (HIV-2) packaging signal is present on all HIV-2 RNA species: cotranslational RNA encapsidation and limitation of Gag protein confer specificity. J Virol *75*, 12058-12069.

Grigorov, B., Decimo, D., Smagulova, F., Pechoux, C., Mougel, M., Muriaux, D., and Darlix, J.L. (2007). Intracellular HIV-1 Gag localization is impaired by mutations in the nucleocapsid zinc fingers. Retrovirology *4*, 54.

Groom, H.C., Anderson, E.C., and Lever, A.M. (2009). Rev: beyond nuclear export. J Gen Virol *90*, 1303-1318.

Guan, Y., Diallo, K., Detorio, M., Whitney, J.B., Liang, C., and Wainberg, M.A. (2001a). Partial restoration of replication of simian immunodeficiency virus by point mutations in either the dimerization initiation site (DIS) or Gag region after deletion mutagenesis within the DIS. J Virol *75*, 11920-11923.

Guan, Y., Whitney, J.B., Diallo, K., and Wainberg, M.A. (2000). Leader sequences downstream of the primer binding site are important for efficient replication of simian immunodeficiency virus. J Virol *74*, 8854-8860.

Guan, Y., Whitney, J.B., Liang, C., and Wainberg, M.A. (2001b). Novel, live attenuated simian immunodeficiency virus constructs containing major deletions in leader RNA sequences. J Virol *75*, 2776-2785.

Guatelli, J. Personal communication.

Guerra-Peraza, O., de Tapia, M., Hohn, T., and Hemmings-Mieszczak, M. (2000). Interaction of the cauliflower mosaic virus coat protein with the pregenomic RNA leader. J Virol *74*, 2067-2072.

Guo, J., Henderson, L.E., Bess, J., Kane, B., and Levin, J.G. (1997). Human immunodeficiency virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-stop DNA. J Virol *71*, 5178-5188.

Guyader, M., Emerman, M., Sonigo, P., Clavel, F., Montagnier, L., and Alizon, M. (1987). Genome organization and transactivation of the human immunodeficiency virus type 2. Nature *326*, 662-669.

Haddrick, M., Lear, A.L., Cann, A.J., and Heaphy, S. (1996). Evidence that a kissing loop structure facilitates genomic RNA dimerisation in HIV-1. J Mol Biol *259*, 58-68.

Hagan, N., and Fabris, D. (2003). Direct mass spectrometric determination of the stoichiometry and binding affinity of the complexes between nucleocapsid protein and RNA stem-loop hairpins of the HIV-1 Psi-recognition element. Biochemistry *42*, 10736-10745.

Hahn, B.H., Kong, L.I., Lee, S.W., Kumar, P., Taylor, M.E., Arya, S.K., and Shaw, G.M. (1987). Relation of HTLV-4 to simian and human immunodeficiency-associated viruses. Nature *330*, 184-186.

Hahn, B.H., Shaw, G.M., De Cock, K.M., and Sharp, P.M. (2000). AIDS as a zoonosis: scientific and public health implications. Science *287*, 607-614.

Han, Y., Wind-Rotolo, M., Yang, H.C., Siliciano, J.D., and Siliciano, R.F. (2007). Experimental approaches to the study of HIV-1 latency. Nat Rev Microbiol *5*, 95-106.

Hargittai, M.R., Gorelick, R.J., Rouzina, I., and Musier-Forsyth, K. (2004). Mechanistic insights into the kinetics of HIV-1 nucleocapsid protein-facilitated tRNA annealing to the primer binding site. J Mol Biol *337*, 951-968.

Harrison, G.P., and Lever, A.M. (1992). The human immunodeficiency virus type 1 packaging signal and major splice donor region have a conserved stable secondary structure. J Virol *66*, 4144-4153.

Harrison, G.P., Miele, G., Hunter, E., and Lever, A.M. (1998). Functional analysis of the core human immunodeficiency virus type 1 packaging signal in a permissive cell line. J Virol *72*, 5886-5896.

Hatziioannou, T., Perez-Caballero, D., Cowan, S., and Bieniasz, P.D. (2005). Cyclophilin interactions with incoming human immunodeficiency virus type 1 capsids with opposing effects on infectivity in human cells. J Virol *79*, 176-183.

Hayashi, T., Shioda, T., Iwakura, Y., and Shibuta, H. (1992). RNA packaging signal of human immunodeficiency virus type 1. Virology *188*, 590-599.

Heinzinger, N.K., Bukinsky, M.I., Haggerty, S.A., Ragland, A.M., Kewalramani, V., Lee, M.A., Gendelman, H.E., Ratner, L., Stevenson, M., and Emerman, M. (1994). The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. Proc Natl Acad Sci USA *91*, 7311-7315.

Helga-Maria, C., Hammarskjold, M.L., and Rekosh, D. (1999). An intact TAR element and cytoplasmic localization are necessary for efficient packaging of human immunodeficiency virus type 1 genomic RNA. J Virol *73*, 4127-4135.

Herbreteau, C.H., Weill, L., Decimo, D., Prevot, D., Darlix, J.L., Sargueil, B., and Ohlmann, T. (2005). HIV-2 genomic RNA contains a novel type of IRES located downstream of its initiation codon. Nat Struct Mol Biol *12*, 1001-1007.

Hibbert, C.S., Mirro, J., and Rein, A. (2004). mRNA molecules containing murine leukemia virus packaging signals are encapsidated as dimers. J Virol *78*, 10927-10938.

Hill, M.K., Shehu-Xhilaga, M., Campbell, S.M., Poumbourios, P., Crowe, S.M., and Mak, J. (2003). The dimer initiation sequence stem-loop of human immunodeficiency virus type 1 is dispensable for viral replication in peripheral blood mononuclear cells. J Virol *77*, 8329-8335.

Hill, M.K., Shehu-Xhilaga, M., Crowe, S.M., and Mak, J. (2002). Proline residues within spacer peptide p1 are important for human immunodeficiency virus type 1 infectivity, protein processing, and genomic RNA dimer stability. J Virol *76*, 11245-11253.

Hirsch, V.M., Olmsted, R.A., Murphey-Corb, M., Purcell, R.H., and Johnson, P.R. (1989). An African primate lentivirus (SIVsm) closely related to HIV-2. Nature *339*, 389-392.

Hockley, D.J., Wood, R.D., Jacobs, J.P., and Garrett, A.J. (1988). Electron microscopy of human immunodeficiency virus. J Gen Virol *69 (Pt 10)*, 2455-2469.

Hogue, I.B., Hoppe, A., and Ono, A. (2009). Quantitative FRET Microscopy Analysis of HIV-1 Gag-Gag Interaction: The Relative Contributions of CA and NC Domains, and Membrane Binding. J Virol, doi: 10.1128/JVI.02545-02508 [Epub ahead of print].

Holmes, R.K., Malim, M.H., and Bishop, K.N. (2007). APOBEC-mediated viral restriction: not simply editing? Trends Biochem Sci *32*, 118-128.

Hooper, E. (1999). The River: A Journey Back to the Source of HIV and AIDS, 1st edn (Penguin, London).

Hu, W.S., and Temin, H.M. (1990a). Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. Proc Natl Acad Sci USA *87*, 1556-1560.

Hu, W.S., and Temin, H.M. (1990b). Retroviral recombination and reverse transcription. Science *250*, 1227-1233.

Huang, M., and Martin, M.A. (1997). Incorporation of Pr160(gag-pol) into virus particles requires the presence of both the major homology region and adjacent C-terminal capsid sequences within the Gag-Pol polyprotein. J Virol *71*, 4472-4478.

Huang, M., Orenstein, J.M., Martin, M.A., and Freed, E.O. (1995). p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. J Virol *69*, 6810-6818.

Huang, Y., Khorchid, A., Gabor, J., Wang, J., Li, X., Darlix, J.L., Wainberg, M.A., and Kleiman, L. (1998). The role of nucleocapsid and U5 stem/A-rich loop sequences in tRNA(3Lys) genomic placement and initiation of reverse transcription in human immunodeficiency virus type 1. J Virol *72*, 3907-3915.

Hubner, W., McNerney, G.P., Chen, P., Dale, B.M., Gordon, R.E., Chuang, F.Y., Li, X.D., Asmuth, D.M., Huser, T., and Chen, B.K. (2009). Quantitative 3D video microscopy of HIV transfer across T cell virological synapses. Science *323*, 1743-1747.

Huet, T., Cheynier, R., Meyerhans, A., Roelants, G., and Wain-Hobson, S. (1990). Genetic organization of a chimpanzee lentivirus related to HIV-1. Nature *345*, 356-359.

Huthoff, H., and Berkhout, B. (2001). Two alternating structures of the HIV-1 leader RNA. RNA *7*, 143-157.

Huthoff, H., and Towers, G.J. (2008). Restriction of retroviral replication by APOBEC3G/F and TRIM5alpha. Trends Microbiol *16*, 612-619.

Inoue, H., Nojima, H., and Okayama, H. (1990). High efficiency transformation of Escherichia coli with plasmids. Gene *96*, 23-28.

Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J., and Varmus, H.E. (1988). Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature *331*, 280-283.

Jacobs, E., Gheysen, D., Thines, D., Francotte, M., and de Wilde, M. (1989). The HIV-1 Gag precursor Pr55gag synthesized in yeast is myristoylated and targeted to the plasma membrane. Gene *79*, 71-81.

Jaffar, S., Grant, A.D., Whitworth, J., Smith, P.G., and Whittle, H. (2004). The natural history of HIV-1 and HIV-2 infections in adults in Africa: a literature review. Bull World Health Organ *82*, 462-469.

Jaffar, S., Wilkins, A., Ngom, P.T., Sabally, S., Corrah, T., Bangali, J.E., Rolfe, M., and Whittle, H.C. (1997). Rate of decline of percentage CD4+ cells is faster in HIV-1 than in HIV-2 infection. J Acquir Immune Defic Syndr Hum Retrovirol *16*, 327-332.

Jakobovits, A., Smith, D.H., Jakobovits, E.B., and Capon, D.J. (1988). A discrete element 3' of human immunodeficiency virus 1 (HIV-1) and HIV-2 mRNA initiation sites mediates transcriptional activation by an HIV trans activator. Mol Cell Biol *8*, 2555-2561.

Jayasena, V.K., Brown, D., Shtatland, T., and Gold, L. (1996). In vitro selection of RNA specifically cleaved by bacteriophage T4 RegB endonuclease. Biochemistry *35*, 2349-2356.

Jenkins, Y., Pornillos, O., Rich, R.L., Myszka, D.G., Sundquist, W.I., and Malim, M.H. (2001). Biochemical analyses of the interactions between human immunodeficiency virus type 1 Vpr and p6(Gag). J Virol *75*, 10537-10542.

Jennes, W., Camara, M., Dieye, T., Mboup, S., and Kestens, L. (2008). Higher homologous and lower cross-reactive Gag-specific T-cell responses in human immunodeficiency virus type 2 (HIV-2) than in HIV-1 infection. J Virol *82*, 8619-8628.

Jones, K.L., Sonza, S., and Mak, J. (2008). Primary T-lymphocytes rescue the replication of HIV-1 DIS RNA mutants in part by facilitating reverse transcription. Nucleic Acids Res *36*, 1578-1588.

Jorgenson, R.L., Vogt, V.M., and Johnson, M.C. (2009). Foreign glycoproteins can be actively recruited to virus assembly sites during pseudotyping. J Virol *83*, 4060-4067.

Joshi, A., Ablan, S.D., Soheilian, F., Nagashima, K., and Freed, E.O. (2009). Evidence that productive human immunodeficiency virus type 1 assembly can occur in an intracellular compartment. J Virol *83*, 5375-5387.

Joshi, A., Garg, H., Nagashima, K., Bonifacino, J.S., and Freed, E.O. (2008). GGA and Arf proteins modulate retrovirus assembly and release. Mol Cell *30*, 227-238.

Joshi, A., Nagashima, K., and Freed, E.O. (2006). Mutation of dileucine-like motifs in the human immunodeficiency virus type 1 capsid disrupts virus assembly, gag-gag interactions, gag-membrane binding, and virion maturation. J Virol *80*, 7939-7951.

Jossinet, F., Lodmell, J.S., Ehresmann, C., Ehresmann, B., and Marquet, R. (2001). Identification of the in vitro HIV-2/SIV RNA dimerization site reveals striking differences with HIV-1. J Biol Chem *276*, 5598-5604.

Jouvenet, N., Bieniasz, P.D., and Simon, S.M. (2008). Imaging the biogenesis of individual HIV-1 virions in live cells. Nature *454*, 236-240.

Jouvenet, N., Neil, S.J., Bess, C., Johnson, M.C., Virgen, C.A., Simon, S.M., and Bieniasz, P.D. (2006). Plasma membrane is the site of productive HIV-1 particle assembly. PLoS Biol *4*, e435.

Jowett, J.B., Hockley, D.J., Nermut, M.V., and Jones, I.M. (1992). Distinct signals in human immunodeficiency virus type 1 Pr55 necessary for RNA binding and particle formation. J Gen Virol *73 (Pt 12)*, 3079-3086.

Kafaie, J., Dolatshahi, M., Ajamian, L., Song, R., Mouland, A.J., Rouiller, I., and Laughrea, M. (2009). Role of capsid sequence and immature nucleocapsid proteins p9 and p15 in Human Immunodeficiency Virus type 1 genomic RNA dimerization. Virology *385*, 233-244.

Kanki, P.J., Alroy, J., and Essex, M. (1985a). Isolation of T-lymphotropic retrovirus related to HTLV-III/LAV from wild-caught African green monkeys. Science *230*, 951-954.

Kanki, P.J., Barin, F., M'Boup, S., Allan, J.S., Romet-Lemonne, J.L., Marlink, R., McLane, M.F., Lee, T.H., Arbeille, B., Denis, F., *et al.* (1986). New human T-lymphotropic retrovirus related to simian T-lymphotropic virus type III (STLV-IIIAGM). Science *232*, 238-243.

Kanki, P.J., Kurth, R., Becker, W., Dreesman, G., McLane, M.F., and Essex, M. (1985b). Antibodies to simian T-lymphotropic retrovirus type III in African green monkeys and recognition of STLV-III viral proteins by AIDS and related sera. Lancet *1*, 1330-1332.

Kanki, P.J., McLane, M.F., King, N.W., Jr., Letvin, N.L., Hunt, R.D., Sehgal, P., Daniel, M.D., Desrosiers, R.C., and Essex, M. (1985c). Serologic identification and

characterization of a macaque T-lymphotropic retrovirus closely related to HTLV-III. Science *228*, 1199-1201.

Kanki, P.J., Travers, K.U., S, M.B., Hsieh, C.C., Marlink, R.G., Gueye, N.A., Siby, T., Thior, I., Hernandez-Avila, M., Sankale, J.L., *et al.* (1994). Slower heterosexual spread of HIV-2 than HIV-1. Lancet *343*, 943-946.

Karacostas, V., Wolffe, E.J., Nagashima, K., Gonda, M.A., and Moss, B. (1993). Overexpression of the HIV-1 gag-pol polyprotein results in intracellular activation of HIV-1 protease and inhibition of assembly and budding of virus-like particles. Virology *193*, 661-671.

Katz, R.A., Terry, R.W., and Skalka, A.M. (1986). A conserved cis-acting sequence in the 5' leader of avian sarcoma virus RNA is required for packaging. J Virol *59*, 163-167.

Kaye, J.F., and Lever, A.M. (1996). trans-acting proteins involved in RNA encapsidation and viral assembly in human immunodeficiency virus type 1. J Virol *70*, 880-886.

Kaye, J.F., and Lever, A.M. (1998). Nonreciprocal packaging of human immunodeficiency virus type 1 and type 2 RNA: a possible role for the p2 domain of Gag in RNA encapsidation. J Virol *72*, 5877-5885.

Kaye, J.F., and Lever, A.M. (1999). Human immunodeficiency virus types 1 and 2 differ in the predominant mechanism used for selection of genomic RNA for encapsidation. J Virol *73*, 3023-3031.

Keele, B.F., Van Heuverswyn, F., Li, Y., Bailes, E., Takehisa, J., Santiago, M.L., Bibollet-Ruche, F., Chen, Y., Wain, L.V., Liegeois, F., *et al.* (2006). Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. Science *313*, 523-526.

Kewalramani, V.N., and Emerman, M. (1996). Vpx association with mature core structures of HIV-2. Virology *218*, 159-168.

Kim, H.J., Lee, K., and O'Rear, J.J. (1994). A short sequence upstream of the 5' major splice site is important for encapsidation of HIV-1 genomic RNA. Virology *198*, 336-340.

Kingsman, S.M., and Kingsman, A.J. (1996). The regulation of human immunodeficiency virus type-1 gene expression. Eur J Biochem *240*, 491-507.

Kirchhoff, F., Greenough, T.C., Brettler, D.B., Sullivan, J.L., and Desrosiers, R.C. (1995). Brief report: absence of intact nef sequences in a long-term survivor with nonprogressive HIV-1 infection. N Engl J Med *332*, 228-232.

Klatzmann, D., Barre-Sinoussi, F., Nugeyre, M.T., Danquet, C., Vilmer, E., Griscelli, C., Brun-Veziret, F., Rouzioux, C., Gluckman, J.C., Chermann, J.C., *et al.* (1984a). Selective tropism of lymphadenopathy associated virus (LAV) for helper-inducer T lymphocytes. Science *225*, 59-63.

Klatzmann, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., Gluckman, J.C., and Montagnier, L. (1984b). T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. Nature *312*, 767-768.

Klein, K.C., Reed, J.C., and Lingappa, J.R. (2007). Intracellular destinies: degradation, targeting, assembly, and endocytosis of HIV Gag. AIDS Rev *9*, 150-161.

Kondo, E., and Gottlinger, H.G. (1996). A conserved LXXLF sequence is the major determinant in p6gag required for the incorporation of human immunodeficiency virus type 1 Vpr. J Virol *70*, 159-164.

Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B.H., Wolinsky, S., and Bhattacharya, T. (2000). Timing the ancestor of the HIV-1 pandemic strains. Science *288*, 1789-1796.

Krausslich, H.G. (2008). A revised model for HIV-1 morphogenesis. Abstract presented at "The retrovirus assembly meeting", October 4-8 2008. (Prague, Czech Republic).

Kung, H.J., Hu, S., Bender, W., Bailey, J.M., Davidson, N., Nicolson, M.O., and McAllister, R.M. (1976). RD-114, baboon, and woolly monkey viral RNA's compared in size and structure. Cell *7*, 609-620.

L'Hernault, A., Greatorex, J.S., Crowther, R.A., and Lever, A.M. (2007). Dimerisation of HIV-2 genomic RNA is linked to efficient RNA packaging, normal particle maturation and viral infectivity. Retrovirology *4*, 90.

Lanchy, J.M., Ivanovitch, J.D., and Lodmell, J.S. (2003a). A structural linkage between the dimerization and encapsidation signals in HIV-2 leader RNA. RNA *9*, 1007-1018.

Lanchy, J.M., and Lodmell, J.S. (2002). Alternate usage of two dimerization initiation sites in HIV-2 viral RNA in vitro. J Mol Biol *319*, 637-648.

Lanchy, J.M., and Lodmell, J.S. (2007). An extended stem-loop 1 is necessary for human immunodeficiency virus type 2 replication and affects genomic RNA encapsidation. J Virol *81*, 3285-3292.

Lanchy, J.M., Rentz, C.A., Ivanovitch, J.D., and Lodmell, J.S. (2003b). Elements located upstream and downstream of the major splice donor site influence the ability of HIV-2 leader RNA to dimerize in vitro. Biochemistry *42*, 2634-2642.

Lanchy, J.M., Szafran, Q.N., and Lodmell, J.S. (2004). Splicing affects presentation of RNA dimerization signals in HIV-2 in vitro. Nucleic Acids Res *32*, 4585-4595.

Langley, C.L., Benga-De, E., Critchlow, C.W., Ndoye, I., Mbengue-Ly, M.D., Kuypers, J., Woto-Gaye, G., Mboup, S., Bergeron, C., Holmes, K.K., *et al.* (1996). HIV-1, HIV-2, human papillomavirus infection and cervical neoplasia in high-risk African women. AIDS *10*, 413-417.

Lassen, K., Han, Y., Zhou, Y., Siliciano, J., and Siliciano, R.F. (2004). The multifactorial nature of HIV-1 latency. Trends Mol Med *10*, 525-531.

Laughrea, M., and Jette, L. (1994). A 19-nucleotide sequence upstream of the 5' major splice donor is part of the dimerization domain of human immunodeficiency virus 1 genomic RNA. Biochemistry *33*, 13464-13474.

Laughrea, M., and Jette, L. (1996). Kissing-loop model of HIV-1 genome dimerization: HIV-1 RNAs can assume alternative dimeric forms, and all sequences upstream or downstream of hairpin 248-271 are dispensable for dimer formation. Biochemistry *35*, 1589-1598.

Laughrea, M., Jette, L., Mak, J., Kleiman, L., Liang, C., and Wainberg, M.A. (1997). Mutations in the kissing-loop hairpin of human immunodeficiency virus type 1 reduce viral infectivity as well as genomic RNA packaging and dimerization. J Virol *71*, 3397-3406.

Laughrea, M., Shen, N., Jette, L., Darlix, J.L., Kleiman, L., and Wainberg, M.A. (2001). Role of distal zinc finger of nucleocapsid protein in genomic RNA dimerization of human immunodeficiency virus type 1; no role for the palindrome crowning the R-U5 hairpin. Virology *281*, 109-116.

Laughrea, M., Shen, N., Jette, L., and Wainberg, M.A. (1999). Variant effects of non-native kissing-loop hairpin palindromes on HIV replication and HIV RNA dimerization: role of stem-loop B in HIV replication and HIV RNA dimerization. Biochemistry *38*, 226-234.

Lawrence, D.C., Stover, C.C., Noznitsky, J., Wu, Z., and Summers, M.F. (2003). Structure of the intact stem and bulge of HIV-1 Psi-RNA stem-loop SL1. J Mol Biol *326*, 529-542.

Le Gall, S., Erdtmann, L., Benichou, S., Berlioz-Torrent, C., Liu, L., Benarous, R., Heard, J.M., and Schwartz, O. (1998). Nef interacts with the mu subunit of clathrin adaptor complexes and reveals a cryptic sorting signal in MHC I molecules. Immunity *8*, 483-495.

Le Rouzic, E., Belaidouni, N., Estrabaud, E., Morel, M., Rain, J.C., Transy, C., and Margottin-Goguet, F. (2007). HIV1 Vpr arrests the cell cycle by recruiting DCAF1/VprBP, a receptor of the Cul4-DDB1 ubiquitin ligase. Cell Cycle *6*, 182-188.

Lee, P.P., and Linial, M.L. (1994). Efficient particle formation can occur if the matrix domain of human immunodeficiency virus type 1 Gag is substituted by a myristylation signal. J Virol *68*, 6644-6654.

Leligdowicz, A., Yindom, L.M., Onyango, C., Sarge-Njie, R., Alabi, A., Cotten, M., Vincent, T., da Costa, C., Aaby, P., Jaye, A., *et al.* (2007). Robust Gag-specific T cell responses characterize viremia control in HIV-2 infection. J Clin Invest *117*, 3067-3074.

Lever, A., Gottlinger, H., Haseltine, W., and Sodroski, J. (1989). Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. J Virol *63*, 4085-4087.

Lever, A., M.,L., . Personal communication.

Lever, A.M. (2007). HIV-1 RNA packaging. Adv Pharmacol *55*, 1-32.

Levesque, K., Halvorsen, M., Abrahamyan, L., Chatel-Chaix, L., Poupon, V., Gordon, H., DesGroseillers, L., Gatignol, A., and Mouland, A.J. (2006). Trafficking of HIV-1 RNA is mediated by heterogeneous nuclear ribonucleoprotein A2 expression and impacts on viral assembly. Traffic *7*, 1177-1193.

Levy, J.A., Hoffman, A.D., Kramer, S.M., Landis, J.A., Shimabukuro, J.M., and Oshiro, L.S. (1984). Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. Science *225*, 840-842.

Li, S., Hill, C.P., Sundquist, W.I., and Finch, J.T. (2000). Image reconstructions of helical assemblies of the HIV-1 CA protein. Nature *407*, 409-413.

Li, X., Quan, Y., Arts, E.J., Li, Z., Preston, B.D., de Rocquigny, H., Roques, B.P., Darlix, J.L., Kleiman, L., Parniak, M.A., *et al.* (1996). Human immunodeficiency virus Type 1 nucleocapsid protein (NCp7) directs specific initiation of minus-strand DNA synthesis primed by human tRNA(Lys3) in vitro: studies of viral RNA molecules mutated in regions that flank the primer binding site. J Virol *70*, 4996-5004.

Liang, C., Rong, L., Cherry, E., Kleiman, L., Laughrea, M., and Wainberg, M.A. (1999a). Deletion mutagenesis within the dimerization initiation site of human immunodeficiency virus type 1 results in delayed processing of the p2 peptide from precursor proteins. J Virol *73*, 6147-6151.

Liang, C., Rong, L., Laughrea, M., Kleiman, L., and Wainberg, M.A. (1998). Compensatory point mutations in the human immunodeficiency virus type 1 Gag region

that are distal from deletion mutations in the dimerization initiation site can restore viral replication. J Virol *72*, 6629-6636.

Liang, C., Rong, L., Quan, Y., Laughrea, M., Kleiman, L., and Wainberg, M.A. (1999b). Mutations within four distinct gag proteins are required to restore replication of human immunodeficiency virus type 1 after deletion mutagenesis within the dimerization initiation site. J Virol *73*, 7014-7020.

Liang, C., Rong, L., Russell, R.S., and Wainberg, M.A. (2000). Deletion Mutagenesis Downstream of the 5' Long Terminal Repeat of Human Immunodeficiency Virus Type 1 Is Compensated for by Point Mutations in both the U5 Region and gag Gene. J Virol *74*, 6251-6261.

Lifson, A.R., Buchbinder, S.P., Sheppard, H.W., Mawle, A.C., Wilber, J.C., Stanley, M., Hart, C.E., Hessol, N.A., and Holmberg, S.D. (1991). Long-term human immunodeficiency virus infection in asymptomatic homosexual and bisexual men with normal CD4+ lymphocyte counts: immunologic and virologic characteristics. J Infect Dis *163*, 959-965.

Lindwasser, O.W., Chaudhuri, R., and Bonifacino, J.S. (2007). Mechanisms of CD4 downregulation by the Nef and Vpu proteins of primate immunodeficiency viruses. Curr Mol Med *7*, 171-184.

Liu, R., Paxton, W.A., Choe, S., Ceradini, D., Martin, S.R., Horuk, R., MacDonald, M.E., Stuhlmann, H., Koup, R.A., and Landau, N.R. (1996). Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. Cell *86*, 367-377.

Lodmell, J.S., Ehresmann, C., Ehresmann, B., and Marquet, R. (2000). Convergence of natural and artificial evolution on an RNA loop-loop interaction: the HIV-1 dimerization initiation site. RNA *6*, 1267-1276.

Lopez-Verges, S., Camus, G., Blot, G., Beauvoir, R., Benarous, R., and Berlioz-Torrent, C. (2006). Tail-interacting protein TIP47 is a connector between Gag and Env and is required for Env incorporation into HIV-1 virions. Proc Natl Acad Sci USA *103*, 14947-14952.

Loussert-Ajaka, I., Chaix, M.L., Korber, B., Letourneur, F., Gomas, E., Allen, E., Ly, T.D., Brun-Vezinet, F., Simon, F., and Saragosti, S. (1995). Variability of human immunodeficiency virus type 1 group O strains isolated from Cameroonian patients living in France. J Virol *69*, 5640-5649.

Lu, Y.L., Bennett, R.P., Wills, J.W., Gorelick, R., and Ratner, L. (1995). A leucine triplet repeat sequence (LXX)4 in p6gag is important for Vpr incorporation into human immunodeficiency virus type 1 particles. J Virol *69*, 6873-6879.

Luban, J., Bossolt, K.L., Franke, E.K., Kalpana, G.V., and Goff, S.P. (1993). Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B. Cell *73*, 1067-1078.

Luban, J., and Goff, S.P. (1994). Mutational analysis of cis-acting packaging signals in human immunodeficiency virus type 1 RNA. J Virol *68*, 3784-3793.

Lusso, P., Cocchi, F., Balotta, C., Markham, P.D., Louie, A., Farci, P., Pal, R., Gallo, R.C., and Reitz, M.S., Jr. (1995). Growth of macrophage-tropic and primary human immunodeficiency virus type 1 (HIV-1) isolates in a unique CD4+ T-cell clone (PM1): failure to downregulate CD4 and to interfere with cell-line-tropic HIV-1. J Virol *69*, 3712-3720.

Lutcke, H. (1995). Signal recognition particle (SRP), a ubiquitous initiator of protein translocation. Eur J Biochem *228*, 531-550.

MacNeil, A., Sankale, J.L., Meloni, S.T., Sarr, A.D., Mboup, S., and Kanki, P. (2006). Genomic sites of human immunodeficiency virus type 2 (HIV-2) integration: similarities to HIV-1 in vitro and possible differences in vivo. J Virol *80*, 7316-7321.

MacNeil, A., Sarr, A.D., Sankale, J.L., Meloni, S.T., Mboup, S., and Kanki, P. (2007). Direct evidence of lower viral replication rates in vivo in human immunodeficiency virus type 2 (HIV-2) infection than in HIV-1 infection. J Virol *81*, 5325-5330.

Mahalingam, S., Ayyavoo, V., Patel, M., Kieber-Emmons, T., and Weiner, D.B. (1997). Nuclear import, virion incorporation, and cell cycle arrest/differentiation are mediated by distinct functional domains of human immunodeficiency virus type 1 Vpr. J Virol *71*, 6339-6347.

Malim, M.H., Bohnlein, S., Fenrick, R., Le, S.Y., Maizel, J.V., and Cullen, B.R. (1989a). Functional comparison of the Rev trans-activators encoded by different primate immunodeficiency virus species. Proc Natl Acad Sci USA *86*, 8222-8226.

Malim, M.H., and Cullen, B.R. (1991). HIV-1 structural gene expression requires the binding of multiple Rev monomers to the viral RRE: implications for HIV-1 latency. Cell *65*, 241-248.

Malim, M.H., and Emerman, M. (2008). HIV-1 accessory proteins--ensuring viral survival in a hostile environment. Cell Host Microbe *3*, 388-398.

Malim, M.H., Hauber, J., Le, S.Y., Maizel, J.V., and Cullen, B.R. (1989b). The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. Nature *338*, 254-257.

Malim, M.H., Tiley, L.S., McCarn, D.F., Rusche, J.R., Hauber, J., and Cullen, B.R. (1990). HIV-1 structural gene expression requires binding of the Rev trans-activator to its RNA target sequence. Cell *60*, 675-683.

Mammano, F., Ohagen, A., Hoglund, S., and Gottlinger, H.G. (1994). Role of the major homology region of human immunodeficiency virus type 1 in virion morphogenesis. J Virol *68*, 4927-4936.

Mann, R., Mulligan, R.C., and Baltimore, D. (1983). Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell *33*, 153-159.

Marcello, A. (2006). Latency: the hidden HIV-1 challenge. Retrovirology *3*, 7.

Mariani, R., Chen, D., Schrofelbauer, B., Navarro, F., Konig, R., Bollman, B., Munk, C., Nymark-McMahon, H., and Landau, N.R. (2003). Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. Cell *114*, 21-31.

Marquet, R., Baudin, F., Gabus, C., Darlix, J.L., Mougel, M., Ehresmann, C., and Ehresmann, B. (1991). Dimerization of human immunodeficiency virus (type 1) RNA: stimulation by cations and possible mechanism. Nucleic Acids Res *19*, 2349-2357.

Marquet, R., Paillart, J.C., Skripkin, E., Ehresmann, C., and Ehresmann, B. (1994). Dimerization of human immunodeficiency virus type 1 RNA involves sequences located upstream of the splice donor site. Nucleic Acids Res *22*, 145-151.

Martin-Serrano, J. (2007). The role of ubiquitin in retroviral egress. Traffic *8*, 1297-1303.

Martin-Serrano, J., Zang, T., and Bieniasz, P.D. (2001). HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. Nat Med *7*, 1313-1319.

Martin, N., and Sattentau, Q. (2009). Cell-to-cell HIV-1 spread and its implications for immune evasion. Curr Opin HIV AIDS *4*, 143-149.

Martinez, N.W., Xue, X., Berro, R.G., Kreitzer, G., and Resh, M.D. (2008). Kinesin KIF4 regulates intracellular trafficking and stability of the human immunodeficiency virus type 1 Gag polyprotein. J Virol *82*, 9937-9950.

Marx, P.A., Li, Y., Lerche, N.W., Sutjipto, S., Gettie, A., Yee, J.A., Brotman, B.H., Prince, A.M., Hanson, A., Webster, R.G., *et al.* (1991). Isolation of a simian

immunodeficiency virus related to human immunodeficiency virus type 2 from a west African pet sooty mangabey. J Virol *65*, 4480-4485.

Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J Mol Biol *288*, 911-940.

Matsui, T., Tanaka, T., Endoh, H., Sato, K., Tanaka, H., Miyauchi, E., Kawashima, Y., Nagai-Makabe, M., Komatsu, H., Kohno, T., *et al.* (2009). The RNA recognition mechanism of human immunodeficiency virus (HIV) type 2 NCp8 is different from that of HIV-1 NCp7. Biochemistry *48*, 4314-4323.

Mavilio, D., Benjamin, J., Daucher, M., Lombardo, G., Kottilil, S., Planta, M.A., Marcenaro, E., Bottino, C., Moretta, L., Moretta, A., *et al.* (2003). Natural killer cells in HIV-1 infection: dichotomous effects of viremia on inhibitory and activating receptors and their functional correlates. Proc Natl Acad Sci USA *100*, 15011-15016.

McBride, M.S., and Panganiban, A.T. (1996). The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures. J Virol *70*, 2963-2973.

McBride, M.S., and Panganiban, A.T. (1997). Position dependence of functional hairpins important for human immunodeficiency virus type 1 RNA encapsidation in vivo. J Virol *71*, 2050-2058.

McBride, M.S., Schwartz, M.D., and Panganiban, A.T. (1997). Efficient encapsidation of human immunodeficiency virus type 1 vectors and further characterization of cis elements required for encapsidation. J Virol *71*, 4544-4554.

McCann, E. (1996). Analysis of the role of the HIV-2 leader sequence in genomic RNA packaging. Ph.D. Department of Medicine: University of Cambridge.

McCann, E.M., and Lever, A.M. (1997). Location of cis-acting signals important for RNA encapsidation in the leader sequence of human immunodeficiency virus type 2. J Virol *71*, 4133-4137.

McCune, J.M., Rabin, L.B., Feinberg, M.B., Lieberman, M., Kosek, J.C., Reyes, G.R., and Weissman, I.L. (1988). Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus. Cell *53*, 55-67.

McKnight, A., Dittmar, M.T., Moniz-Periera, J., Ariyoshi, K., Reeves, J.D., Hibbitts, S., Whitby, D., Aarons, E., Proudfoot, A.E., Whittle, H., *et al.* (1998). A broad range of chemokine receptors are used by primary isolates of human immunodeficiency virus type 2 as coreceptors with CD4. J Virol *72*, 4065-4071.

McLaughlin, S., and Murray, D. (2005). Plasma membrane phosphoinositide organization by protein electrostatics. Nature *438*, 605-611.

Meric, C., and Goff, S.P. (1989). Characterization of Moloney murine leukemia virus mutants with single-amino-acid substitutions in the Cys-His box of the nucleocapsid protein. J Virol *63*, 1558-1568.

Meric, C., Gouilloud, E., and Spahr, P.F. (1988). Mutations in Rous sarcoma virus nucleocapsid protein p12 (NC): deletions of Cys-His boxes. J Virol *62*, 3328-3333.

Meric, C., and Spahr, P.F. (1986). Rous sarcoma virus nucleic acid-binding protein p12 is necessary for viral 70S RNA dimer formation and packaging. J Virol *60*, 450-459.

Miele, G., Mouland, A., Harrison, G.P., Cohen, E., and Lever, A.M. (1996). The human immunodeficiency virus type 1 5' packaging signal structure affects translation but does not function as an internal ribosome entry site structure. J Virol *70*, 944-951.

Mikkelsen, J.G., Lund, A.H., Duch, M., and Pedersen, F.S. (2000). Mutations of the kissing-loop dimerization sequence influence the site specificity of murine leukemia virus recombination in vivo. J Virol *74*, 600-610.

Mikkelsen, J.G., Rasmussen, S.V., and Pedersen, F.S. (2004). Complementarity-directed RNA dimer-linkage promotes retroviral recombination in vivo. Nucleic Acids Res *32*, 102-114.

Mildvan, D., Mathur, U., Enlow, R.W., Romain, P.L., Winchester, R.J., Colp, C., Singman, H., Adelsberg, B.R., and Spigland, I. (1982). Opportunistic infections and immune deficiency in homosexual men. Ann Intern Med *96*, 700-704.

Mitchell, R.S., Beitzel, B.F., Schroder, A.R., Shinn, P., Chen, H., Berry, C.C., Ecker, J.R., and Bushman, F.D. (2004). Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol *2*, E234.

Mok, H.P., and Lever, A. (2008). Waking Up the Sleepers: HIV Latency and Reactivation. J Formos Med Assoc *107*, 909-914.

Moore, M.D., Fu, W., Nikolaitchik, O., Chen, J., Ptak, R.G., and Hu, W.S. (2007). Dimer initiation signal of human immunodeficiency virus type 1: its role in partner selection during RNA copackaging and its effects on recombination. J Virol *81*, 4002-4011.

Morita, E., and Sundquist, W.I. (2004). Retrovirus budding. Annu Rev Cell Dev Biol *20*, 395-425.

Mortlock, D., Keller, E.B., Ziegra, C.J., and Suter, M.M. (1993). High efficiency transfection of monkey COS-1 cells. J Tissue Culture Methods *15*, 176-180.

Mouland, A.J., Mercier, J., Luo, M., Bernier, L., DesGroseillers, L., and Cohen, E.A. (2000). The double-stranded RNA-binding protein Staufen is incorporated in human immunodeficiency virus type 1: evidence for a role in genomic RNA encapsidation. J Virol *74*, 5441-5451.

Mouland, A.J., Xu, H., Cui, H., Krueger, W., Munro, T.P., Prasol, M., Mercier, J., Rekosh, D., Smith, R., Barbarese, E., *et al.* (2001). RNA trafficking signals in human immunodeficiency virus type 1. Mol Cell Biol *21*, 2133-2143.

Mueller, S.M., Jung, R., Weiler, S., and Lang, S.M. (2004). Vpx proteins of SIVmac239 and HIV-2 ROD interact with the cytoskeletal protein alpha-actinin 1. J Gen Virol *85*, 3291-3303.

Mujeeb, A., Clever, J.L., Billeci, T.M., James, T.L., and Parslow, T.G. (1998). Structure of the dimer initiation complex of HIV-1 genomic RNA. Nat Struct Biol *5*, 432-436.

Mujeeb, A., Parslow, T.G., Zarrinpar, A., Das, C., and James, T.L. (1999). NMR structure of the mature dimer initiation complex of HIV-1 genomic RNA. FEBS Lett *458*, 387-392.

Mujeeb, A., Ulyanov, N.B., Georgantis, S., Smirnov, I., Chung, J., Parslow, T.G., and James, T.L. (2007). Nucleocapsid protein-mediated maturation of dimer initiation complex of full-length SL1 stemloop of HIV-1: sequence effects and mechanism of RNA refolding. Nucleic Acids Res *35*, 2026-2034.

Murakami, T. (2008). Roles of the interactions between Env and Gag proteins in the HIV-1 replication cycle. Microbiol Immunol *52*, 287-295.

Murakami, T., and Freed, E.O. (2000). Genetic evidence for an interaction between human immunodeficiency virus type 1 matrix and alpha-helix 2 of the gp41 cytoplasmic tail. J Virol *74*, 3548-3554.

Muriaux, D., De Rocquigny, H., Roques, B.P., and Paoletti, J. (1996a). NCp7 activates HIV-1Lai RNA dimerization by converting a transient loop- loop complex into a stable dimer. J Biol Chem *271*, 33686-33692.

Muriaux, D., Fosse, P., and Paoletti, J. (1996b). A kissing complex together with a stable dimer is involved in the HIV- 1Lai RNA dimerization process in vitro. Biochemistry *35*, 5075-5082.

Muriaux, D., Girard, P.M., Bonnet-Mathoniere, B., and Paoletti, J. (1995). Dimerization of HIV-1Lai RNA at low ionic strength. An autocomplementary sequence in the 5' leader region is evidenced by an antisense oligonucleotide. J Biol Chem *270*, 8209-8216.

Muriaux, D., Mirro, J., Harvin, D., and Rein, A. (2001). RNA is a structural element in retrovirus particles. Proc Natl Acad Sci USA *98*, 5246-5251.

Muriaux, D., Mirro, J., Nagashima, K., Harvin, D., and Rein, A. (2002). Murine leukemia virus nucleocapsid mutant particles lacking viral RNA encapsidate ribosomes. J Virol *76*, 11405-11413.

Myers, E.L., and Allen, J.F. (2002). Tsg101, an inactive homologue of ubiquitin ligase e2, interacts specifically with human immunodeficiency virus type 2 gag polyprotein and results in increased levels of ubiquitinated gag. J Virol *76*, 11226-11235.

Neil, S.J. Personal communication.

Neil, S.J., Sandrin, V., Sundquist, W.I., and Bieniasz, P.D. (2007). An interferon-alpha-induced tethering mechanism inhibits HIV-1 and Ebola virus particle release but is counteracted by the HIV-1 Vpu protein. Cell Host Microbe *2*, 193-203.

Neil, S.J., Zang, T., and Bieniasz, P.D. (2008). Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature *451*, 425-430.

Nermut, M.V., Hockley, D.J., Jowett, J.B., Jones, I.M., Garreau, M., and Thomas, D. (1994). Fullerene-like organization of HIV gag-protein shell in virus-like particles produced by recombinant baculovirus. Virology *198*, 288-296.

Nishi, M., Ryo, A., Tsurutani, N., Ohba, K., Sawasaki, T., Morishita, R., Perrem, K., Aoki, I., Morikawa, Y., and Yamamoto, N. (2009). Requirement for microtubule integrity in the SOCS1-mediated intracellular dynamics of HIV-1 Gag. FEBS Lett *583*, 1243-1250.

Norrgren, H., Marquina, S., Leitner, T., Aaby, P., Melbye, M., Poulsen, A.G., Larsen, O., Dias, F., Escanilla, D., Andersson, S., *et al.* (1997). HIV-2 genetic variation and DNA load in asymptomatic carriers and AIDS cases in Guinea-Bissau. J Acquir Immune Defic Syndr Hum Retrovirol *16*, 31-38.

Nuvor, S.V., van der Sande, M., Rowland-Jones, S., Whittle, H., and Jaye, A. (2006). Natural killer cell function is well preserved in asymptomatic human immunodeficiency virus type 2 (HIV-2) infection but similar to that of HIV-1 infection when CD4 T-cell counts fall. J Virol *80*, 2529-2538.

Nydegger, S., Foti, M., Derdowski, A., Spearman, P., and Thali, M. (2003). HIV-1 egress is gated through late endosomal membranes. Traffic *4*, 902-910.

O'Connor, G.M., Holmes, A., Mulcahy, F., and Gardiner, C.M. (2007). Natural Killer cells from long-term non-progressor HIV patients are characterized by altered phenotype and function. Clin Immunol *124*, 277-283.

Oertle, S., and Spahr, P.F. (1990). Role of the gag polyprotein precursor in packaging and maturation of Rous sarcoma virus genomic RNA. J Virol *64*, 5757-5763.

Ono, A., Ablan, S.D., Lockett, S.J., Nagashima, K., and Freed, E.O. (2004). Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. Proc Natl Acad Sci USA *101*, 14889-14894.

Ono, A., and Freed, E.O. (1999). Binding of human immunodeficiency virus type 1 Gag to membrane: role of the matrix amino terminus. J Virol *73*, 4136-4144.

Ono, A., and Freed, E.O. (2004). Cell-type-dependent targeting of human immunodeficiency virus type 1 assembly to the plasma membrane and the multivesicular body. J Virol *78*, 1552-1563.

Ono, A., Huang, M., and Freed, E.O. (1997). Characterization of human immunodeficiency virus type 1 matrix revertants: effects on virus assembly, Gag processing, and Env incorporation into virions. J Virol *71*, 4409-4418.

Ono, A., Orenstein, J.M., and Freed, E.O. (2000). Role of the Gag matrix domain in targeting human immunodeficiency virus type 1 assembly. J Virol *74*, 2855-2866.

Ooms, M., Huthoff, H., Russell, R., Liang, C., and Berkhout, B. (2004). A riboswitch regulates RNA dimerization and packaging in human immunodeficiency virus type 1 virions. J Virol *78*, 10814-10819.

Paillart, J.C., Berthoux, L., Ottmann, M., Darlix, J.L., Marquet, R., Ehresmann, B., and Ehresmann, C. (1996a). A dual role of the putative RNA dimerization initiation site of human immunodeficiency virus type 1 in genomic RNA packaging and proviral DNA synthesis. J Virol *70*, 8348-8354.

Paillart, J.C., Dettenhofer, M., Yu, X.F., Ehresmann, C., Ehresmann, B., and Marquet, R. (2004a). First snapshots of the HIV-1 RNA structure in infected cells and in virions. J Biol Chem *279*, 48397-48403.

Paillart, J.C., and Gottlinger, H.G. (1999). Opposing effects of human immunodeficiency virus type 1 matrix mutations support a myristyl switch model of gag membrane targeting. J Virol *73*, 2604-2612.

Paillart, J.C., Marquet, R., Skripkin, E., Ehresmann, B., and Ehresmann, C. (1994). Mutational analysis of the bipartite dimer linkage structure of human immunodeficiency virus type 1 genomic RNA. J Biol Chem *269*, 27486-27493.

Paillart, J.C., Shehu-Xhilaga, M., Marquet, R., and Mak, J. (2004b). Dimerization of retroviral RNA genomes:an inseparable pair. Nat Rev Microbiol *2*, 461-472.

Paillart, J.C., Skripkin, E., Ehresmann, B., Ehresmann, C., and Marquet, R. (1996b). A loop-loop "kissing" complex is the essential part of the dimer linkage of genomic HIV-1 RNA. Proc Natl Acad Sci USA *93*, 5572-5577.

Paillart, J.C., Skripkin, E., Ehresmann, B., Ehresmann, C., and Marquet, R. (2002). In vitro evidence for a long range pseudoknot in the 5'-untranslated and matrix coding regions of HIV-1 genomic RNA. J Biol Chem *277*, 5995-6004.

Paillart, J.C., Westhof, E., Ehresmann, C., Ehresmann, B., and Marquet, R. (1997). Non-canonical interactions in a kissing loop complex: the dimerization initiation site of HIV-1 genomic RNA. J Mol Biol *270*, 36-49.

Palmer, E., Martin, M.L., Goldsmith, C., and Switzer, W. (1988). Ultrastructure of human immunodeficiency virus type 2. J Gen Virol *69 (Pt 6)*, 1425-1429.

Pancio, H.A., and Ratner, L. (1998). Human immunodeficiency virus type 2 Vpx-Gag interaction. J Virol *72*, 5271-5275.

Pancio, H.A., Vander Heyden, N., and Ratner, L. (2000). The C-terminal proline-rich tail of human immunodeficiency virus type 2 Vpx is necessary for nuclear localization of the viral preintegration complex in nondividing cells. J Virol *74*, 6162-6167.

Panganiban, A.T., and Fiore, D. (1988). Ordered interstrand and intrastrand DNA transfer during reverse transcription. Science *241*, 1064-1069.

Pantaleo, G., Demarest, J.F., Schacker, T., Vaccarezza, M., Cohen, O.J., Daucher, M., Graziosi, C., Schnittman, S.S., Quinn, T.C., Shaw, G.M., *et al.* (1997). The qualitative nature of the primary immune response to HIV infection is a prognosticator of disease progression independent of the initial level of plasma viremia. Proc Natl Acad Sci USA *94*, 254-258.

Pantaleo, G., and Fauci, A.S. (1996). Immunopathogenesis of HIV infection. Annu Rev Microbiol *50*, 825-854.

Pantaleo, G., Graziosi, C., Demarest, J.F., Butini, L., Montroni, M., Fox, C.H., Orenstein, J.M., Kotler, D.P., and Fauci, A.S. (1993a). HIV infection is active and progressive in lymphoid tissue during the clinically latent stage of disease. Nature *362*, 355-358.

Pantaleo, G., Graziosi, C., and Fauci, A.S. (1993b). New concepts in the immunopathogenesis of human immunodeficiency virus infection. N Engl J Med *328*, 327-335.

Pantaleo, G., Menzo, S., Vaccarezza, M., Graziosi, C., Cohen, O.J., Demarest, J.F., Montefiori, D., Orenstein, J.M., Fox, C., Schrager, L.K., *et al.* (1995). Studies in

subjects with long-term nonprogressive human immunodeficiency virus infection. N Engl J Med *332*, 209-216.

Parent, L.J., Cairns, T.M., Albert, J.A., Wilson, C.B., Wills, J.W., and Craven, R.C. (2000). RNA dimerization defect in a Rous sarcoma virus matrix mutant. J Virol *74*, 164-172.

Parolin, C., Dorfman, T., Palu, G., Gottlinger, H., and Sodroski, J. (1994). Analysis in human immunodeficiency virus type 1 vectors of cis-acting sequences that affect gene transfer into human lymphocytes. J Virol *68*, 3888-3895.

Patel, J., Wang, S.W., Izmailova, E., and Aldovini, A. (2003). The simian immunodeficiency virus 5' untranslated leader sequence plays a role in intracellular viral protein accumulation and in RNA packaging. J Virol *77*, 6284-6292.

Paxton, W., Connor, R.I., and Landau, N.R. (1993). Incorporation of Vpr into human immunodeficiency virus type 1 virions: requirement for the p6 region of gag and mutational analysis. J Virol *67*, 7229-7237.

Peeters, M., Gueye, A., Mboup, S., Bibollet-Ruche, F., Ekaza, E., Mulanga, C., Ouedrago, R., Gandji, R., Mpele, P., Dibanga, G., *et al.* (1997). Geographical distribution of HIV-1 group O viruses in Africa. AIDS *11*, 493-498.

Peeters, M., Honore, C., Huet, T., Bedjabaga, L., Ossari, S., Bussi, P., Cooper, R.W., and Delaporte, E. (1989). Isolation and partial characterization of an HIV-related virus occurring naturally in chimpanzees in Gabon. AIDS *3*, 625-630.

Pelchen-Matthews, A., Kramer, B., and Marsh, M. (2003). Infectious HIV-1 assembles in late endosomes in primary macrophages. J Cell Biol *162*, 443-455.

Perlman, M., and Resh, M.D. (2006). Identification of an intracellular trafficking and assembly pathway for HIV-1 gag. Traffic *7*, 731-745.

Pettit, S.C., Moody, M.D., Wehbie, R.S., Kaplan, A.H., Nantermet, P.V., Klein, C.A., and Swanstrom, R. (1994). The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions. J Virol *68*, 8017-8027.

Piatak, M., Jr., Saag, M.S., Yang, L.C., Clark, S.J., Kappes, J.C., Luk, K.C., Hahn, B.H., Shaw, G.M., and Lifson, J.D. (1993). High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. Science *259*, 1749-1754.

Poeschla, E., Gilbert, J., Li, X., Huang, S., Ho, A., and Wong-Staal, F. (1998). Identification of a human immunodeficiency virus type 2 (HIV-2) encapsidation

determinant and transduction of nondividing human cells by HIV-2-based lentivirus vectors. J Virol *72*, 6527-6536.

Poole, E., Strappe, P., Mok, H.P., Hicks, R., and Lever, A.M. (2005). HIV-1 Gag-RNA interaction occurs at a perinuclear/centrosomal site; analysis by confocal microscopy and FRET. Traffic *6*, 741-755.

Poon, D.T., Chertova, E.N., and Ott, D.E. (2002). Human immunodeficiency virus type 1 preferentially encapsidates genomic RNAs that encode Pr55(Gag): functional linkage between translation and RNA packaging. Virology *293*, 368-378.

Poon, D.T., Li, G., and Aldovini, A. (1998). Nucleocapsid and matrix protein contributions to selective human immunodeficiency virus type 1 genomic RNA packaging. J Virol *72*, 1983-1993.

Poon, D.T., Wu, J., and Aldovini, A. (1996). Charged amino acid residues of human immunodeficiency virus type 1 nucleocapsid p7 protein involved in RNA packaging and infectivity. J Virol *70*, 6607-6616.

Popovic, M., Sarngadharan, M.G., Read, E., and Gallo, R.C. (1984). Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. Science *224*, 497-500.

Popper, S.J., Sarr, A.D., Travers, K.U., Gueye-Ndiaye, A., Mboup, S., Essex, M.E., and Kanki, P.J. (1999). Lower human immunodeficiency virus (HIV) type 2 viral load reflects the difference in pathogenicity of HIV-1 and HIV-2. J Infect Dis *180*, 1116-1121.

Post, K., Kankia, B., Gopalakrishnan, S., Yang, V., Cramer, E., Saladores, P., Gorelick, R.J., Guo, J., Musier-Forsyth, K., and Levin, J.G. (2009). Fidelity of plus-strand priming requires the nucleic acid chaperone activity of HIV-1 nucleocapsid protein. Nucleic Acids Res *37*, 1755-1766.

Potts, B. (1990). Mini RT assay. In Techniques in HIV research, A. Aldovini, and B. Walker, eds. (New York, Stockton Press), pp. 103-106.

Poulsen, A.G., Kvinesdal, B., Aaby, P., Molbak, K., Frederiksen, K., Dias, F., and Lauritzen, E. (1989). Prevalence of and mortality from human immunodeficiency virus type 2 in Bissau, West Africa. Lancet *1*, 827-831.

Prats, A.C., Roy, C., Wang, P.A., Erard, M., Housset, V., Gabus, C., Paoletti, C., and Darlix, J.L. (1990). Cis elements and trans-acting factors involved in dimer formation of murine leukemia virus RNA. J Virol *64*, 774-783.

Prats, A.C., Sarih, L., Gabus, C., Litvak, S., Keith, G., and Darlix, J.L. (1988). Small finger protein of avian and murine retroviruses has nucleic acid annealing activity and positions the replication primer tRNA onto genomic RNA. EMBO J *7*, 1777-1783.

Quinones-Mateu, M.E., Ball, S.C., Marozsan, A.J., Torre, V.S., Albright, J.L., Vanham, G., van Der Groen, G., Colebunders, R.L., and Arts, E.J. (2000). A dual infection/competition assay shows a correlation between ex vivo human immunodeficiency virus type 1 fitness and disease progression. J Virol *74*, 9222-9233.

Raposo, G., Moore, M., Innes, D., Leijendekker, R., Leigh-Brown, A., Benaroch, P., and Geuze, H. (2002). Human macrophages accumulate HIV-1 particles in MHC II compartments. Traffic *3*, 718-729.

Reeves, J.D., and Doms, R.W. (2002). Human immunodeficiency virus type 2. J Gen Virol *83*, 1253-1265.

Reeves, J.D., McKnight, A., Potempa, S., Simmons, G., Gray, P.W., Power, C.A., Wells, T., Weiss, R.A., and Talbot, S.J. (1997). CD4-independent infection by HIV-2 (ROD/B): use of the 7-transmembrane receptors CXCR-4, CCR-3, and V28 for entry. Virology *231*, 130-134.

Rein, A., Harvin, D.P., Mirro, J., Ernst, S.M., and Gorelick, R.J. (1994). Evidence that a central domain of nucleocapsid protein is required for RNA packaging in murine leukemia virus. J Virol *68*, 6124-6129.

Rein, A., McClure, M.R., Rice, N.R., Luftig, R.B., and Schultz, A.M. (1986). Myristylation site in Pr65gag is essential for virus particle formation by Moloney murine leukemia virus. Proc Natl Acad Sci USA *83*, 7246-7250.

Remy, G. (1998). HIV-2 infection throughout the world. A geographical perspective. Santé *8*, 440-446.

Rhodes, D.I., Ashton, L., Solomon, A., Carr, A., Cooper, D., Kaldor, J., and Deacon, N. (2000). Characterization of three nef-defective human immunodeficiency virus type 1 strains associated with long-term nonprogression. Australian Long-Term Nonprogressor Study Group. J Virol *74*, 10581-10588.

Ricci, E.P., Herbreteau, C.H., Decimo, D., Schaupp, A., Datta, S.A., Rein, A., Darlix, J.L., and Ohlmann, T. (2008). In vitro expression of the HIV-2 genomic RNA is controlled by three distinct internal ribosome entry segments that are regulated by the HIV protease and the Gag polyprotein. RNA *14*, 1443-1455.

Ritter, G.D., Jr., Yamshchikov, G., Cohen, S.J., and Mulligan, M.J. (1996). Human immunodeficiency virus type 2 glycoprotein enhancement of particle budding: role of the cytoplasmic domain. J Virol *70*, 2669-2673.

Rizvi, T.A., and Panganiban, A.T. (1993). Simian immunodeficiency virus RNA is efficiently encapsidated by human immunodeficiency virus type 1 particles. J Virol *67*, 2681-2688.

Rodriguez, S.K., Sarr, A.D., MacNeil, A., Thakore-Meloni, S., Gueye-Ndiaye, A., Traore, I., Dia, M.C., Mboup, S., and Kanki, P.J. (2007). Comparison of heterologous neutralizing antibody responses of human immunodeficiency virus type 1 (HIV-1)- and HIV-2-infected Senegalese patients: distinct patterns of breadth and magnitude distinguish HIV-1 and HIV-2 infections. J Virol *81*, 5331-5338.

Roeth, J.F., and Collins, K.L. (2006). Human immunodeficiency virus type 1 Nef: adapting to intracellular trafficking pathways. Microbiol Mol Biol Rev *70*, 548-563.

Rong, L., Russell, R.S., Hu, J., Guan, Y., Kleiman, L., Liang, C., and Wainberg, M.A. (2001). Hydrophobic amino acids in the human immunodeficiency virus type 1 p2 and nucleocapsid proteins can contribute to the rescue of deleted viral RNA packaging signals. J Virol *75*, 7230-7243.

Rong, L., Russell, R.S., Hu, J., Laughrea, M., Wainberg, M.A., and Liang, C. (2003). Deletion of stem-loop 3 is compensated by second-site mutations within the Gag protein of human immunodeficiency virus type 1. Virology *314*, 221-228.

Rowland-Jones, S.L. (1999). HIV: The deadly passenger in dendritic cells. Curr Biol *9*, R248-250.

Rowland-Jones, S.L., and Whittle, H.C. (2007). Out of Africa: what can we learn from HIV-2 about protective immunity to HIV-1? Nat Immunol *8*, 329-331.

Rulli, S.J., Jr., Hibbert, C.S., Mirro, J., Pederson, T., Biswal, S., and Rein, A. (2007). Selective and Nonselective Packaging of Cellular RNAs in Retrovirus Particles. J Virol *81*, 6623-6631.

Russell, R.S., Hu, J., Beriault, V., Mouland, A.J., Kleiman, L., Wainberg, M.A., and Liang, C. (2003a). Sequences downstream of the 5' splice donor site are required for both packaging and dimerization of human immunodeficiency virus type 1 RNA. J Virol *77*, 84-96.

Russell, R.S., Liang, C., and Wainberg, M.A. (2004). Is HIV-1 RNA dimerization a prerequisite for packaging? Yes, no, probably? Retrovirology *1*, 23.

Russell, R.S., Roldan, A., Detorio, M., Hu, J., Wainberg, M.A., and Liang, C. (2003b). Effects of a single amino acid substitution within the p2 region of human immunodeficiency virus type 1 on packaging of spliced viral RNA. J Virol *77*, 12986-12995.

Ryan-Graham, M.A., and Peden, K.W. (1995). Both virus and host components are important for the manifestation of a Nef- phenotype in HIV-1 and HIV-2. Virology *213*, 158-168.

Ryo, A., Tsurutani, N., Ohba, K., Kimura, R., Komano, J., Nishi, M., Soeda, H., Hattori, S., Perrem, K., Yamamoto, M., *et al.* (2008). SOCS1 is an inducible host factor during HIV-1 infection and regulates the intracellular trafficking and stability of HIV-1 Gag. Proc Natl Acad Sci USA *105*, 294-299.

Saad, J.S., Miller, J., Tai, J., Kim, A., Ghanam, R.H., and Summers, M.F. (2006). Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. Proc Natl Acad Sci USA *103*, 11364-11369.

Sakaguchi, K., Zambrano, N., Baldwin, E.T., Shapiro, B.A., Erickson, J.W., Omichinski, J.G., Clore, G.M., Gronenborn, A.M., and Appella, E. (1993). Identification of a binding site for the human immunodeficiency virus type 1 nucleocapsid protein. Proc Natl Acad Sci USA *90*, 5219-5223.

Sakamoto, T., Morita, S., Tabata, K., Nakamura, K., and Kawai, G. (2002). Solution structure of a SRP19 binding domain in human SRP RNA. J Biochem *132*, 177-182.

Sakuragi, J., Shioda, T., and Panganiban, A.T. (2001). Duplication of the primary encapsidation and dimer linkage region of human immunodeficiency virus type 1 RNA results in the appearance of monomeric RNA in virions. J Virol *75*, 2557-2565.

Sakuragi, J., Ueda, S., Iwamoto, A., and Shioda, T. (2003). Possible role of dimerization in human immunodeficiency virus type 1 genome RNA packaging. J Virol *77*, 4060-4069.

Sakuragi, J.I., and Panganiban, A.T. (1997). Human immunodeficiency virus type 1 RNA outside the primary encapsidation and dimer linkage region affects RNA dimer stability in vivo. J Virol *71*, 3250-3254.

Sambrook, J., and Russell, D.W. (2001). Molecular cloning a laboratory manual, 3rd edn (Cold Spring Harbor (N.Y.), Cold Spring Harbor Laboratory Press).

Samson, M., Libert, F., Doranz, B.J., Rucker, J., Liesnard, C., Farber, C.M., Saragosti, S., Lapoumeroulie, C., Cognaux, J., Forceille, C., *et al.* (1996). Resistance to HIV-1

infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. Nature *382*, 722-725.

Sanford, J.R., Coutinho, P., Hackett, J.A., Wang, X., Ranahan, W., and Caceres, J.F. (2008). Identification of nuclear and cytoplasmic mRNA targets for the shuttling protein SF2/ASF. PLoS ONE *3*, e3369.

Sanford, J.R., Wang, X., Mort, M., Vanduyn, N., Cooper, D.N., Mooney, S.D., Edenberg, H.J., and Liu, Y. (2009). Splicing factor SFRS1 recognizes a functionally diverse landscape of RNA transcripts. Genome Res *19*, 381-394.

Sanson, B., and Uzan, M. (1995). Post-transcriptional controls in bacteriophage T4: roles of the sequence-specific endoribonuclease RegB. FEMS Microbiol Rev *17*, 141-150.

Santiago, M.L., Range, F., Keele, B.F., Li, Y., Bailes, E., Bibollet-Ruche, F., Fruteau, C., Noe, R., Peeters, M., Brookfield, J.F., *et al.* (2005). Simian immunodeficiency virus infection in free-ranging sooty mangabeys (Cercocebus atys atys) from the Tai Forest, Cote d'Ivoire: implications for the origin of epidemic human immunodeficiency virus type 2. J Virol *79*, 12515-12527.

Saphire, A.C., Bobardt, M.D., and Gallay, P.A. (2002). Cyclophilin A plays distinct roles in human immunodeficiency virus type 1 entry and postentry events, as revealed by spinoculation. J Virol *76*, 4671-4677.

Sarngadharan, M.G., Popovic, M., Bruch, L., Schupbach, J., and Gallo, R.C. (1984). Antibodies reactive with human T-lymphotropic retroviruses (HTLV-III) in the serum of patients with AIDS. Science *224*, 506-508.

Schim van der Loeff, M.F., and Aaby, P. (1999). Towards a better understanding of the epidemiology of HIV-2. AIDS *13 Suppl A*, S69-84.

Schindler, M., Munch, J., Kutsch, O., Li, H., Santiago, M.L., Bibollet-Ruche, F., Muller-Trutwin, M.C., Novembre, F.J., Peeters, M., Courgnaud, V., *et al.* (2006). Nef-mediated suppression of T cell activation was lost in a lentiviral lineage that gave rise to HIV-1. Cell *125*, 1055-1067.

Schmalzbauer, E., Strack, B., Dannull, J., Guehmann, S., and Moelling, K. (1996). Mutations of basic amino acids of NCp7 of human immunodeficiency virus type 1 affect RNA binding in vitro. J Virol *70*, 771-777.

Scholz, I., Arvidson, B., Huseby, D., and Barklis, E. (2005). Virus particle core defects caused by mutations in the human immunodeficiency virus capsid N-terminal domain. J Virol *79*, 1470-1479.

Schroder, A.R., Shinn, P., Chen, H., Berry, C., Ecker, J.R., and Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. Cell *110*, 521-529.

Schwartz, M.D., Fiore, D., and Panganiban, A.T. (1997). Distinct functions and requirements for the Cys-His boxes of the human immunodeficiency virus type 1 nucleocapsid protein during RNA encapsidation and replication. J Virol *71*, 9295-9305.

Sfakianos, J.N., LaCasse, R.A., and Hunter, E. (2003). The M-PMV cytoplasmic targeting-retention signal directs nascent Gag polypeptides to a pericentriolar region of the cell. Traffic *4*, 660-670.

Sharova, N., Wu, Y., Zhu, X., Stranska, R., Kaushik, R., Sharkey, M., and Stevenson, M. (2008). Primate lentiviral Vpx commandeers DDB1 to counteract a macrophage restriction. PLoS Pathog *4*, e1000057.

Sharp, P.M., Bailes, E., Chaudhuri, R.R., Rodenburg, C.M., Santiago, M.O., and Hahn, B.H. (2001). The origins of acquired immune deficiency syndrome viruses: where and when? Philos Trans R Soc Lond B Biol Sci *356*, 867-876.

Sharp, P.M., Bailes, E., Gao, F., Beer, B.E., Hirsch, V.M., and Hahn, B.H. (2000). Origins and evolution of AIDS viruses: estimating the time-scale. Biochem Soc Trans *28*, 275-282.

Sharp, P.M., Robertson, D.L., and Hahn, B.H. (1995). Cross-species transmission and recombination of 'AIDS' viruses. Philos Trans R Soc Lond B Biol Sci *349*, 41-47.

Sheehy, A.M., Gaddis, N.C., Choi, J.D., and Malim, M.H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature *418*, 646-650.

Sheehy, A.M., Gaddis, N.C., and Malim, M.H. (2003). The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. Nat Med *9*, 1404-1407.

Shehu-Xhilaga, M., Crowe, S.M., and Mak, J. (2001a). Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. J Virol *75*, 1834-1841.

Shehu-Xhilaga, M., Hill, M., Marshall, J.A., Kappes, J., Crowe, S.M., and Mak, J. (2002). The conformation of the mature dimeric human immunodeficiency virus type 1 RNA genome requires packaging of pol protein. J Virol *76*, 4331-4340.

Shehu-Xhilaga, M., Kraeusslich, H.G., Pettit, S., Swanstrom, R., Lee, J.Y., Marshall, J.A., Crowe, S.M., and Mak, J. (2001b). Proteolytic processing of the p2/nucleocapsid

cleavage site is critical for human immunodeficiency virus type 1 RNA dimer maturation. J Virol *75*, 9156-9164.

Shen, N., Jette, L., Liang, C., Wainberg, M.A., and Laughrea, M. (2000). Impact of human immunodeficiency virus type 1 RNA dimerization on viral infectivity and of stem-loop B on RNA dimerization and reverse transcription and dissociation of dimerization from packaging. J Virol *74*, 5729-5735.

Sheng, N., and Erickson-Viitanen, S. (1994). Cleavage of p15 protein in vitro by human immunodeficiency virus type 1 protease is RNA dependent. J Virol *68*, 6207-6214.

Sheppard, H.W., Lang, W., Ascher, M.S., Vittinghoff, E., and Winkelstein, W. (1993). The characterization of non-progressors: long-term HIV-1 infection with stable CD4+ T-cell levels. AIDS *7*, 1159-1166.

Sherer, N.M., Lehmann, M.J., Jimenez-Soto, L.F., Ingmundson, A., Horner, S.M., Cicchetti, G., Allen, P.G., Pypaert, M., Cunningham, J.M., and Mothes, W. (2003). Visualization of retroviral replication in living cells reveals budding into multivesicular bodies. Traffic *4*, 785-801.

Sherry, B., Zybarth, G., Alfano, M., Dubrovsky, L., Mitchell, R., Rich, D., Ulrich, P., Bucala, R., Cerami, A., and Bukrinsky, M. (1998). Role of cyclophilin A in the uptake of HIV-1 by macrophages and T lymphocytes. Proc Natl Acad Sci USA *95*, 1758-1763.

Shubsda, M.F., Paoletti, A.C., Hudson, B.S., and Borer, P.N. (2002). Affinities of packaging domain loops in HIV-1 RNA for the nucleocapsid protein. Biochemistry *41*, 5276-5282.

Silvestri, G., Sodora, D.L., Koup, R.A., Paiardini, M., O'Neil, S.P., McClure, H.M., Staprans, S.I., and Feinberg, M.B. (2003). Nonpathogenic SIV infection of sooty mangabeys is characterized by limited bystander immunopathology despite chronic high-level viremia. Immunity *18*, 441-452.

Simon, F., Mauclere, P., Roques, P., Loussert-Ajaka, I., Muller-Trutwin, M.C., Saragosti, S., Georges-Courbot, M.C., Barre-Sinoussi, F., and Brun-Vezinet, F. (1998). Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. Nat Med *4*, 1032-1037.

Skripkin, E., Paillart, J.C., Marquet, R., Ehresmann, B., and Ehresmann, C. (1994). Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. Proc Natl Acad Sci USA *91*, 4945-4949.

Smallman-Raynor, M., and Cliff, A. (1991). The spread of human immunodeficiency virus type 2 into Europe: a geographical analysis. Int J Epidemiol *20*, 480-489.

Smith, S.D., Shatsky, M., Cohen, P.S., Warnke, R., Link, M.P., and Glader, B.E. (1984). Monoclonal antibody and enzymatic profiles of human malignant T-lymphoid cells and derived cell lines. Cancer Res *44*, 5657-5660.

Song, H., Nakayama, E.E., Yokoyama, M., Sato, H., Levy, J.A., and Shioda, T. (2007a). A single amino acid of the human immunodeficiency virus type 2 capsid affects its replication in the presence of cynomolgus monkey and human TRIM5alphas. J Virol *81*, 7280-7285.

Song, R., Kafaie, J., Yang, L., and Laughrea, M. (2007b). HIV-1 viral RNA is selected in the form of monomers that dimerize in a three-step protease-dependent process; the DIS of stem-loop 1 initiates viral RNA dimerization. J Mol Biol *371*, 1084-1098.

Sorge, J., Ricci, W., and Hughes, S.H. (1983). cis-Acting RNA packaging locus in the 115-nucleotide direct repeat of Rous sarcoma virus. J Virol *48*, 667-675.

Soriano, V., Gomes, P., Heneine, W., Holguin, A., Doruana, M., Antunes, R., Mansinho, K., Switzer, W.M., Araujo, C., Shanmugam, V., *et al.* (2000). Human immunodeficiency virus type 2 (HIV-2) in Portugal: clinical spectrum, circulating subtypes, virus isolation, and plasma viral load. J Med Virol *61*, 111-116.

South, T.L., Blake, P.R., Sowder, R.C., 3rd, Arthur, L.O., Henderson, L.E., and Summers, M.F. (1990). The nucleocapsid protein isolated from HIV-1 particles binds zinc and forms retroviral-type zinc fingers. Biochemistry *29*, 7786-7789.

Spearman, P., Wang, J.J., Vander Heyden, N., and Ratner, L. (1994). Identification of human immunodeficiency virus type 1 Gag protein domains essential to membrane binding and particle assembly. J Virol *68*, 3232-3242.

Spriggs, S., Garyu, L., Connor, R., and Summers, M.F. (2008). Potential intra- and intermolecular interactions involving the unique-5' region of the HIV-1 5'-UTR. Biochemistry *47*, 13064-13073.

Srinivasakumar, N., Hammarskjold, M.L., and Rekosh, D. (1995). Characterization of deletion mutations in the capsid region of human immunodeficiency virus type 1 that affect particle formation and Gag-Pol precursor incorporation. J Virol *69*, 6106-6114.

Stahl, R.E., Friedman-Kien, A., Dubin, R., Marmor, M., and Zolla-Pazner, S. (1982). Immunologic abnormalities in homosexual men. Relationship to Kaposi's sarcoma. Am J Med *73*, 171-178.

Stewart, L., Schatz, G., and Vogt, V.M. (1990). Properties of avian retrovirus particles defective in viral protease. J Virol *64*, 5076-5092.

Strack, B., Calistri, A., Craig, S., Popova, E., and Gottlinger, H.G. (2003). AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. Cell *114*, 689-699.

Strappe, P.M., Greatorex, J., Thomas, J., Biswas, P., McCann, E., and Lever, A.M. (2003). The packaging signal of simian immunodeficiency virus is upstream of the major splice donor at a distance from the RNA cap site similar to that of human immunodeficiency virus types 1 and 2. J Gen Virol *84*, 2423-2430.

Strappe, P.M., Hampton, D.W., Brown, D., Gonzales, B.C., Caldwell, M., Fawcett, J.W., and Lever, A.M. (2005). Identification of unique reciprocal and non reciprocal cross packaging relationships between HIV-1, HIV-2 and SIV reveals an efficient SIV/HIV-2 lentiviral vector system with highly favourable features for in vivo testing and clinical usage. Retrovirology *2*, 55.

Strebel, K., Klimkait, T., Maldarelli, F., and Martin, M.A. (1989). Molecular and biochemical analyses of human immunodeficiency virus type 1 vpu protein. J Virol *63*, 3784-3791.

Swanson, C.M., and Malim, M.H. (2006). Retrovirus RNA trafficking: from chromatin to invasive genomes. Traffic *7*, 1440-1450.

Switzer, W.M., Wiktor, S., Soriano, V., Silva-Graca, A., Mansinho, K., Coulibaly, I.M., Ekpini, E., Greenberg, A.E., Folks, T.M., and Heneine, W. (1998). Evidence of Nef truncation in human immunodeficiency virus type 2 infection. J Infect Dis *177*, 65-71.

Takahashi, K., Baba, S., Koyanagi, Y., Yamamoto, N., Takaku, H., and Kawai, G. (2001). Two basic regions of NCp7 are sufficient for conformational conversion of HIV-1 dimerization initiation site from kissing-loop dimer to extended-duplex dimer. J Biol Chem *276*, 31274-31278.

Takehisa, J., Kraus, M.H., Ayouba, A., Bailes, E., Van Heuverswyn, F., Decker, J.M., Li, Y., Rudicell, R.S., Learn, G.H., Neel, C., *et al.* (2009). Origin and biology of simian immunodeficiency virus in wild-living western gorillas. J Virol *83*, 1635-1648.

Tanchou, V., Gabus, C., Rogemond, V., and Darlix, J.L. (1995). Formation of stable and functional HIV-1 nucleoprotein complexes in vitro. J Mol Biol *252*, 563-571.

Tang, C., Loeliger, E., Luncsford, P., Kinde, I., Beckett, D., and Summers, M.F. (2004). Entropic switch regulates myristate exposure in the HIV-1 matrix protein. Proc Natl Acad Sci USA *101*, 517-522.

Tang, S., Ablan, S., Dueck, M., Ayala-Lopez, W., Soto, B., Caplan, M., Nagashima, K., Hewlett, I.K., Freed, E.O., and Levin, J.G. (2007). A second-site suppressor

significantly improves the defective phenotype imposed by mutation of an aromatic residue in the N-terminal domain of the HIV-1 capsid protein. Virology *359*, 105-115.

Temin, H.M. (1991). Sex and recombination in retroviruses. Trends Genet *7*, 71-74.

Terwilliger, E.F., Cohen, E.A., Lu, Y.C., Sodroski, J.G., and Haseltine, W.A. (1989). Functional role of human immunodeficiency virus type 1 vpu. Proc Natl Acad Sci USA *86*, 5163-5167.

Thali, M., Bukovsky, A., Kondo, E., Rosenwirth, B., Walsh, C.T., Sodroski, J., and Gottlinger, H.G. (1994). Functional association of cyclophilin A with HIV-1 virions. Nature *372*, 363-365.

Thomas, J.A., and Gorelick, R.J. (2008). Nucleocapsid protein function in early infection processes. Virus Res *134*, 39-63.

Tian, H., and Kole, R. (2001). Strong RNA splicing enhancers identified by a modified method of cycled selection interact with SR protein. J Biol Chem *276*, 33833-33839.

Tiley, L.S., Malim, M.H., Tewary, H.K., Stockley, P.G., and Cullen, B.R. (1992). Identification of a high-affinity RNA-binding site for the human immunodeficiency virus type 1 Rev protein. Proc Natl Acad Sci USA *89*, 758-762.

Torrent, C., Gabus, C., and Darlix, J.L. (1994). A small and efficient dimerization/packaging signal of rat VL30 RNA and its use in murine leukemia virus-VL30-derived vectors for gene transfer. J Virol *68*, 661-667.

Trible, R.P., Emert-Sedlak, L., Wales, T.E., Ayyavoo, V., Engen, J.R., and Smithgall, T.E. (2007). Allosteric loss-of-function mutations in HIV-1 Nef from a long-term non-progressor. J Mol Biol *374*, 121-129.

Tristem, M., Marshall, C., Karpas, A., and Hill, F. (1992). Evolution of the primate lentiviruses: evidence from vpx and vpr. EMBO J *11*, 3405-3412.

Tritch, R.J., Cheng, Y.E., Yin, F.H., and Erickson-Viitanen, S. (1991). Mutagenesis of protease cleavage sites in the human immunodeficiency virus type 1 gag polyprotein. J Virol *65*, 922-930.

Trkola, A., Ketas, T., Kewalramani, V.N., Endorf, F., Binley, J.M., Katinger, H., Robinson, J., Littman, D.R., and Moore, J.P. (1998). Neutralization sensitivity of human immunodeficiency virus type 1 primary isolates to antibodies and CD4-based reagents is independent of coreceptor usage. J Virol *72*, 1876-1885.

UNAIDS, and World Health Organization (2008). Report on the global HIV/AIDS epidemic 2008 (Geneva, UNAIDS).

Usami, Y., Popov, S., and Gottlinger, H.G. (2007). Potent rescue of human immunodeficiency virus type 1 late domain mutants by ALIX/AIP1 depends on its CHMP4 binding site. J Virol *81*, 6614-6622.

Van Damme, N., Goff, D., Katsura, C., Jorgenson, R.L., Mitchell, R., Johnson, M.C., Stephens, E.B., and Guatelli, J. (2008). The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. Cell Host Microbe *3*, 245-252.

van Heuverswyn, F., Li, Y., Neel, C., Bailes, E., Keele, B.F., Liu, W., Loul, S., Butel, C., Liegeois, F., Bienvenue, Y., *et al.* (2006). Human immunodeficiency viruses: SIV infection in wild gorillas. Nature *444*, 164.

van Wamel, J.L., and Berkhout, B. (1998). The first strand transfer during HIV-1 reverse transcription can occur either intramolecularly or intermolecularly. Virology *244*, 245-251.

Vandegraaff, N., and Engelman, A. (2007). Molecular mechanisms of HIV integration and therapeutic intervention. Expert Rev Mol Med *9*, 1-19.

Vanden Haesevelde, M., Decourt, J.L., De Leys, R.J., Vanderborght, B., van der Groen, G., van Heuverswijn, H., and Saman, E. (1994). Genomic cloning and complete sequence analysis of a highly divergent African human immunodeficiency virus isolate. J Virol *68*, 1586-1596.

Varthakavi, V., Heimann-Nichols, E., Smith, R.M., Sun, Y., Bram, R.J., Ali, S., Rose, J., Ding, L., and Spearman, P. (2008). Identification of calcium-modulating cyclophilin ligand as a human host restriction to HIV-1 release overcome by Vpu. Nat Med *14*, 641-647.

von Schwedler, U., Kornbluth, R.S., and Trono, D. (1994). The nuclear localization signal of the matrix protein of human immunodeficiency virus type 1 allows the establishment of infection in macrophages and quiescent T lymphocytes. Proc Natl Acad Sci USA *91*, 6992-6996.

von Schwedler, U.K., Stemmler, T.L., Klishko, V.Y., Li, S., Albertine, K.H., Davis, D.R., and Sundquist, W.I. (1998). Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. EMBO J *17*, 1555-1568.

von Schwedler, U.K., Stray, K.M., Garrus, J.E., and Sundquist, W.I. (2003a). Functional surfaces of the human immunodeficiency virus type 1 capsid protein. J Virol *77*, 5439-5450.

von Schwedler, U.K., Stuchell, M., Muller, B., Ward, D.M., Chung, H.Y., Morita, E., Wang, H.E., Davis, T., He, G.P., Cimbora, D.M., *et al.* (2003b). The protein network of HIV budding. Cell *114*, 701-713.

Wang, C.T., and Barklis, E. (1993). Assembly, processing, and infectivity of human immunodeficiency virus type 1 gag mutants. J Virol *67*, 4264-4273.

Wang, D., Yu, Q.C., Schroer, J., Murphy, E., and Shenk, T. (2007). Human cytomegalovirus uses two distinct pathways to enter retinal pigmented epithelial cells. Proc Natl Acad Sci USA *104*, 20037-20042.

Wang, S.W., and Aldovini, A. (2002). RNA incorporation is critical for retroviral particle integrity after cell membrane assembly of Gag complexes. J Virol *76*, 11853-11865.

Wang, S.W., Noonan, K., and Aldovini, A. (2004). Nucleocapsid-RNA interactions are essential to structural stability but not to assembly of retroviruses. J Virol *78*, 716-723.

Watanabe, S., and Temin, H.M. (1982). Encapsidation sequences for spleen necrosis virus, an avian retrovirus, are between the 5' long terminal repeat and the start of the gag gene. Proc Natl Acad Sci USA *79*, 5986-5990.

Wei, P., Garber, M.E., Fang, S.M., Fischer, W.H., and Jones, K.A. (1998). A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. Cell *92*, 451-462.

Weiss, A., Wiskocil, R.L., and Stobo, J.D. (1984). The role of T3 surface molecules in the activation of human T cells: a two-stimulus requirement for IL 2 production reflects events occurring at a pre-translational level. J Immunol *133*, 123-128.

Welker, R., Hohenberg, H., Tessmer, U., Huckhagel, C., and Krausslich, H.G. (2000). Biochemical and structural analysis of isolated mature cores of human immunodeficiency virus type 1. J Virol *74*, 1168-1177.

Welsch, S., Habermann, A., Jager, S., Muller, B., Krijnse-Locker, J., and Krausslich, H.G. (2006). Ultrastructural analysis of ESCRT proteins suggests a role for endosome-associated tubular-vesicular membranes in ESCRT function. Traffic *7*, 1551-1566.

Wertheim, J.O., and Worobey, M. (2009). Dating the age of the SIV lineages that gave rise to HIV-1 and HIV-2. PLoS Comput Biol *5*, e1000377.

Whitney, J.B., Oliveira, M., Detorio, M., Guan, Y., and Wainberg, M.A. (2002). The M184V mutation in reverse transcriptase can delay reversion of attenuated variants of simian immunodeficiency virus. J Virol *76*, 8958-8962.

Whitney, J.B., and Wainberg, M.A. (2006). Impaired RNA incorporation and dimerization in live attenuated leader-variants of SIVmac239. Retrovirology *3*, 96.

Wickham, L., Duchaine, T., Luo, M., Nabi, I.R., and DesGroseillers, L. (1999). Mammalian staufen is a double-stranded-RNA- and tubulin-binding protein which localizes to the rough endoplasmic reticulum. Mol Cell Biol *19*, 2220-2230.

Wilk, T., Gross, I., Gowen, B.E., Rutten, T., de Haas, F., Welker, R., Krausslich, H.G., Boulanger, P., and Fuller, S.D. (2001). Organization of immature human immunodeficiency virus type 1. J Virol *75*, 759-771.

Wilkins, A., Ricard, D., Todd, J., Whittle, H., Dias, F., and Paulo Da Silva, A. (1993). The epidemiology of HIV infection in a rural area of Guinea-Bissau. AIDS *7*, 1119-1122.

Wilkinson, K.A., Gorelick, R.J., Vasa, S.M., Guex, N., Rein, A., Mathews, D.H., Giddings, M.C., and Weeks, K.M. (2008). High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. PLoS Biol *6*, e96.

Willey, R.L., Maldarelli, F., Martin, M.A., and Strebel, K. (1992). Human immunodeficiency virus type 1 Vpu protein regulates the formation of intracellular gp160-CD4 complexes. J Virol *66*, 226-234.

Wills, J.W., Cameron, C.E., Wilson, C.B., Xiang, Y., Bennett, R.P., and Leis, J. (1994). An assembly domain of the Rous sarcoma virus Gag protein required late in budding. J Virol *68*, 6605-6618.

Wilson, S.A., Sieiro-Vazquez, C., Edwards, N.J., Iourin, O., Byles, E.D., Kotsopoulou, E., Adamson, C.S., Kingsman, S.M., Kingsman, A.J., and Martin-Rendon, E. (1999). Cloning and characterization of hIF2, a human homologue of bacterial translation initiation factor 2, and its interaction with HIV-1 matrix. Biochem J *342 (Pt 1)*, 97-103.

Worobey, M., Gemmel, M., Teuwen, D.E., Haselkorn, T., Kunstman, K., Bunce, M., Muyembe, J.J., Kabongo, J.M., Kalengayi, R.M., Van Marck, E., *et al.* (2008). Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature *455*, 661-664.

Yamaguchi, J., Coffey, R., Vallari, A., Ngansop, C., Mbanya, D., Ndembi, N., Kaptue, L., Gurtler, L.G., Bodelle, P., Schochetman, G., *et al.* (2006). Identification of HIV type 1 group N infections in a husband and wife in Cameroon: viral genome sequences provide evidence for horizontal transmission. AIDS Res Hum Retroviruses *22*, 83-92.

Yamaguchi, J., Devare, S.G., and Brennan, C.A. (2000). Identification of a new HIV-2 subtype based on phylogenetic analysis of full-length genomic sequence. AIDS Res Hum Retroviruses *16*, 925-930.

Yedavalli, V.S., Neuveut, C., Chi, Y.H., Kleiman, L., and Jeang, K.T. (2004). Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function. Cell *119*, 381-392.

Yee, C., Krishnan-Hewlett, I., Baker, C.C., Schlegel, R., and Howley, P.M. (1985). Presence and expression of human papillomavirus sequences in human cervical carcinoma cell lines. Am J Pathol *119*, 361-366.

Yilmaz, A., Bolinger, C., and Boris-Lawrie, K. (2006). Retrovirus translation initiation: Issues and hypotheses derived from study of HIV-1. Curr HIV Res *4*, 131-139.

Ylinen, L.M., Keckesova, Z., Wilson, S.J., Ranasinghe, S., and Towers, G.J. (2005). Differential restriction of human immunodeficiency virus type 2 and simian immunodeficiency virus SIVmac by TRIM5alpha alleles. J Virol *79*, 11580-11587.

Yu, E.T., Hawkins, A., Eaton, J., and Fabris, D. (2008). MS3D structural elucidation of the HIV-1 packaging signal. Proc Natl Acad Sci USA *105*, 12248-12253.

Yu, X., Yuan, X., Matsuda, Z., Lee, T.H., and Essex, M. (1992). The matrix protein of human immunodeficiency virus type 1 is required for incorporation of viral envelope protein into mature virions. J Virol *66*, 4966-4971.

Yu, X.F., Ito, S., Essex, M., and Lee, T.H. (1988). A naturally immunogenic virion-associated protein specific for HIV-2 and SIV. Nature *335*, 262-265.

Zeffman, A., Hassard, S., Varani, G., and Lever, A. (2000). The major HIV-1 packaging signal is an extended bulged stem loop whose structure is altered on interaction with the Gag polyprotein. J Mol Biol *297*, 877-893.

Zhang, W.H., Hockley, D.J., Nermut, M.V., Morikawa, Y., and Jones, I.M. (1996). Gag-Gag interactions in the C-terminal domain of human immunodeficiency virus type 1 p24 capsid antigen are essential for Gag particle assembly. J Gen Virol *77*, 743-751.

Zhang, Y., and Barklis, E. (1995). Nucleocapsid protein effects on the specificity of retrovirus RNA encapsidation. J Virol *69*, 5716-5722.

Zhou, W., Parent, L.J., Wills, J.W., and Resh, M.D. (1994). Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. J Virol *68*, 2556-2569.

Zhu, H., Jian, H., and Zhao, L.J. (2004). Identification of the 15FRFG domain in HIV-1 Gag p6 essential for Vpr packaging into the virion. Retrovirology *1*, 26.

Zhu, T., Korber, B.T., Nahmias, A.J., Hooper, E., Sharp, P.M., and Ho, D.D. (1998). An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. Nature *391*, 594-597.

Zhu, Y., Pe'ery, T., Peng, J., Ramanathan, Y., Marshall, N., Marshall, T., Amendt, B., Mathews, M.B., and Price, D.H. (1997). Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro. Genes Dev *11*, 2622-2632.

Zimmerman, C., Klein, K.C., Kiser, P.K., Singh, A.R., Firestein, B.L., Riba, S.C., and Lingappa, J.R. (2002). Identification of a host protein essential for assembly of immature HIV-1 capsids. Nature *415*, 88-92.

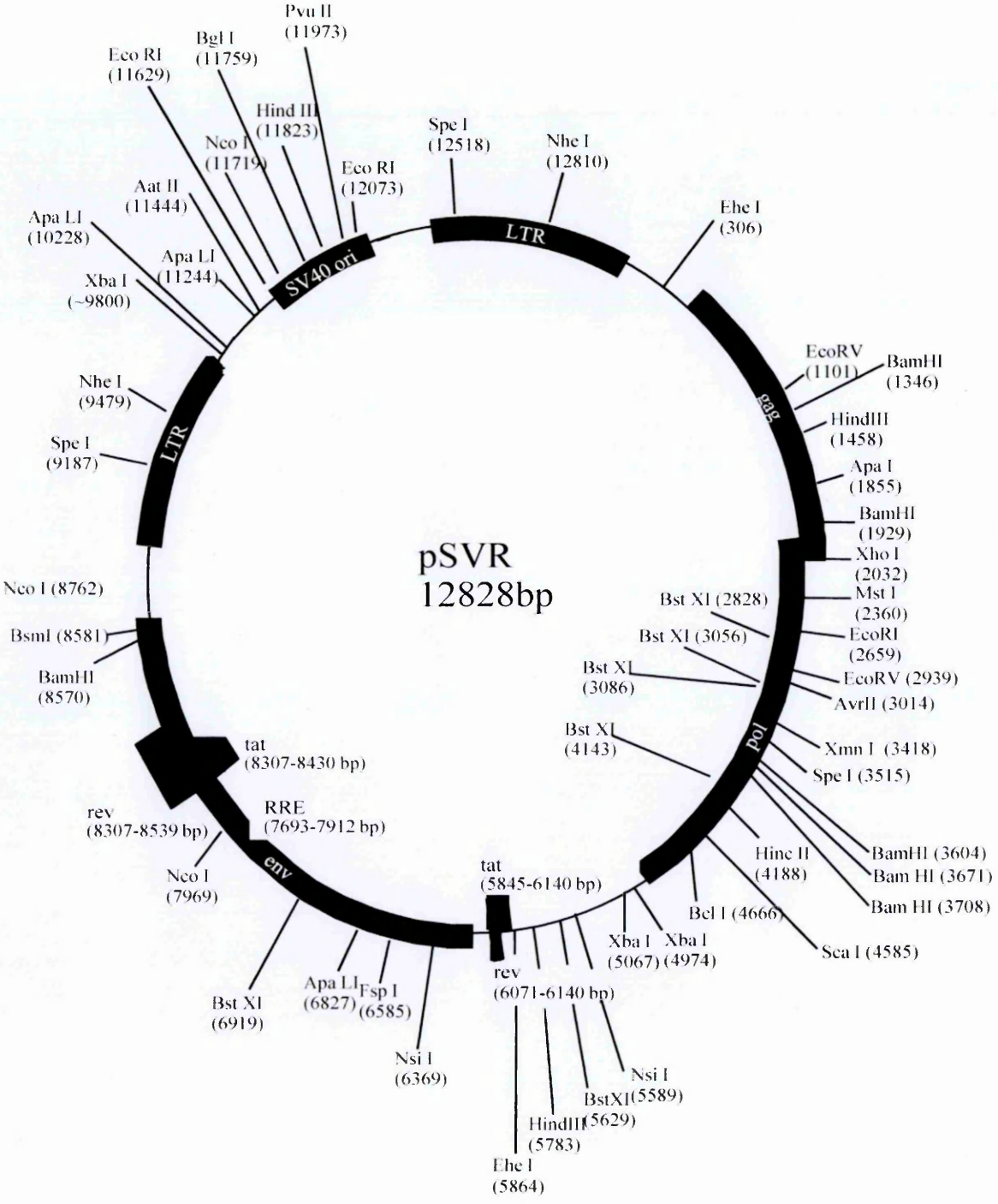
Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res *31*, 3406-3415.

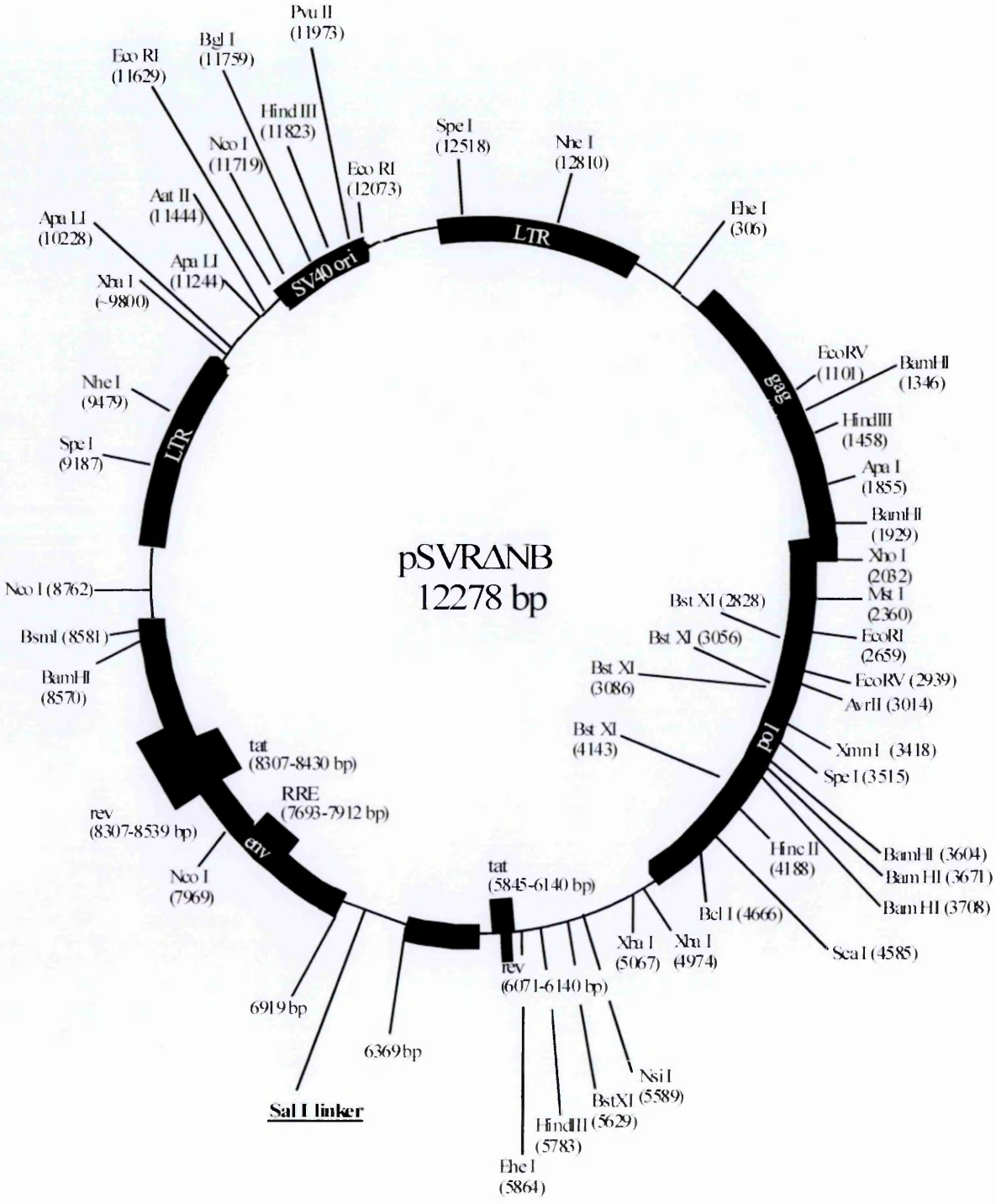
Appendix 1: Publications derived from this study

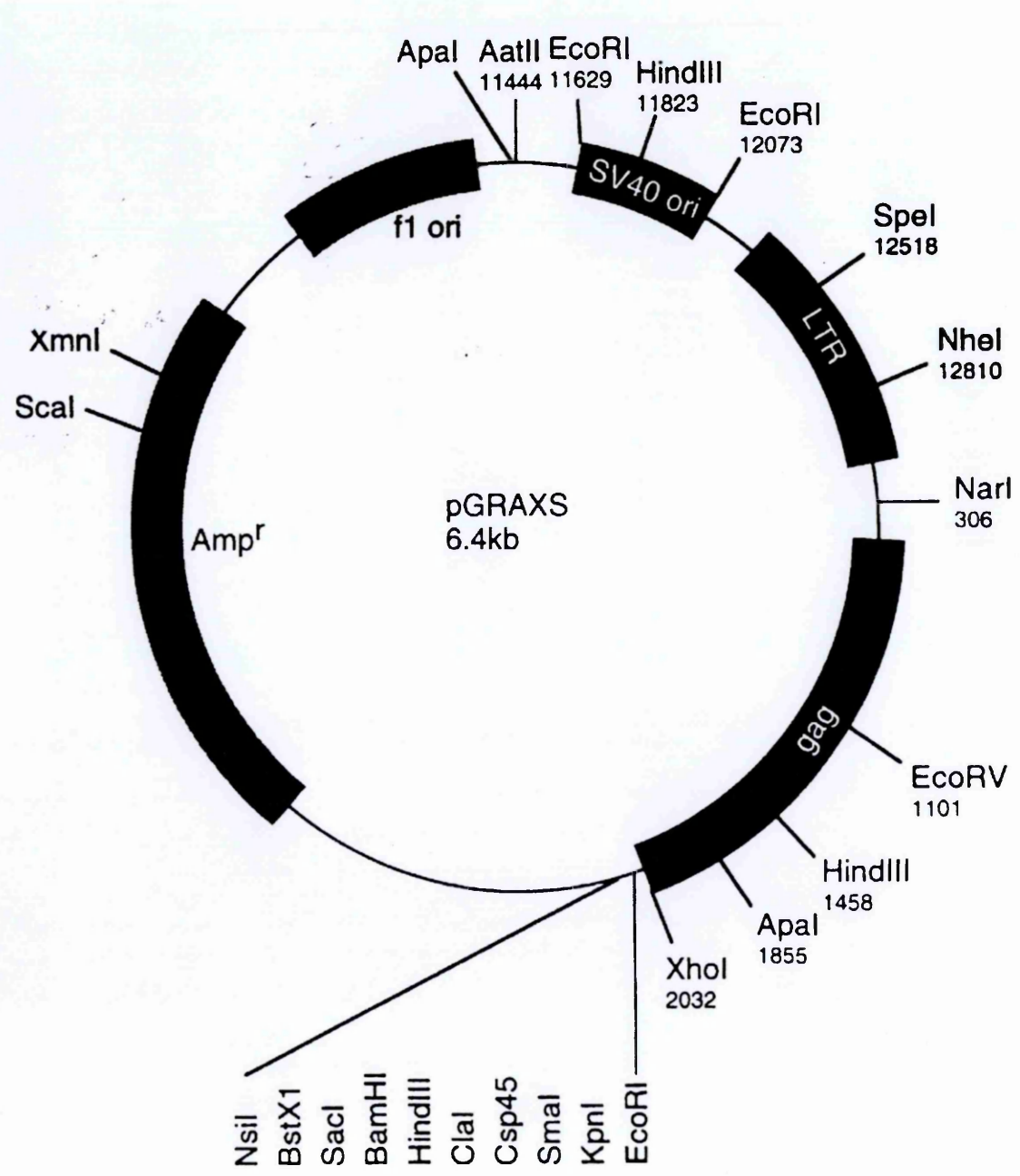
L'Hernault A., Greatorex J.S., Crowther R.A. and Lever A.M. (2007). Dimerisation of HIV-2 genomic RNA is linked to efficient RNA packaging, normal particle maturation and viral infectivity. Retrovirology 4: 90.

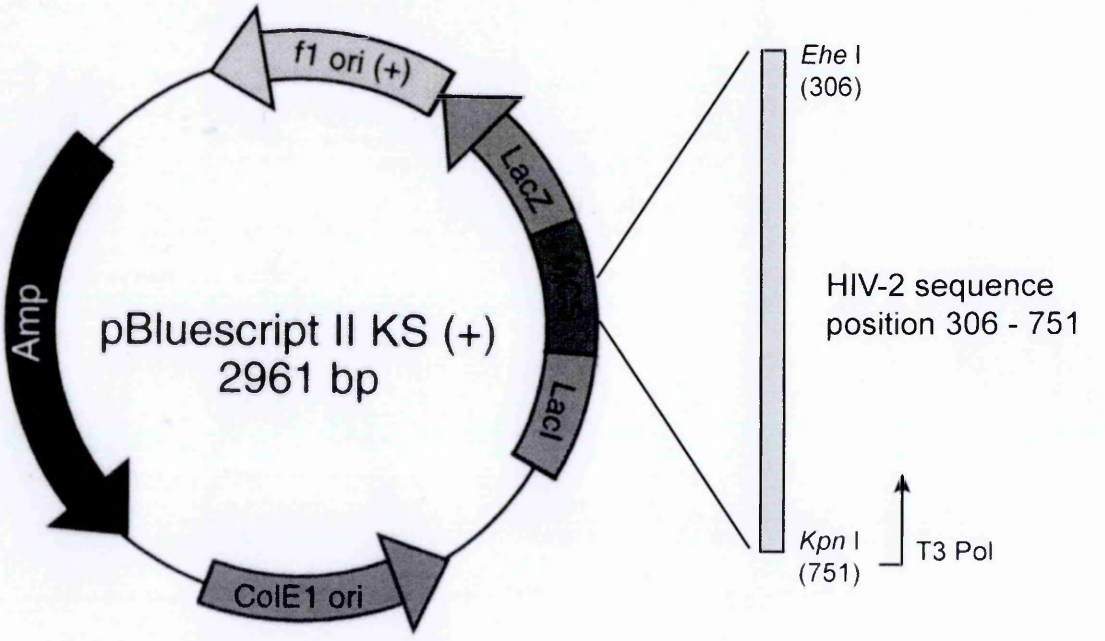
L'Hernault A., Li W., Greatorex J.S. and Lever A.M. (2009). Identification of a GGAG motif in the HIV-2 RNA leader that is essential for genome dimerisation and viral replication. *Manuscript in preparation.*

Appendix 2: Plasmid maps

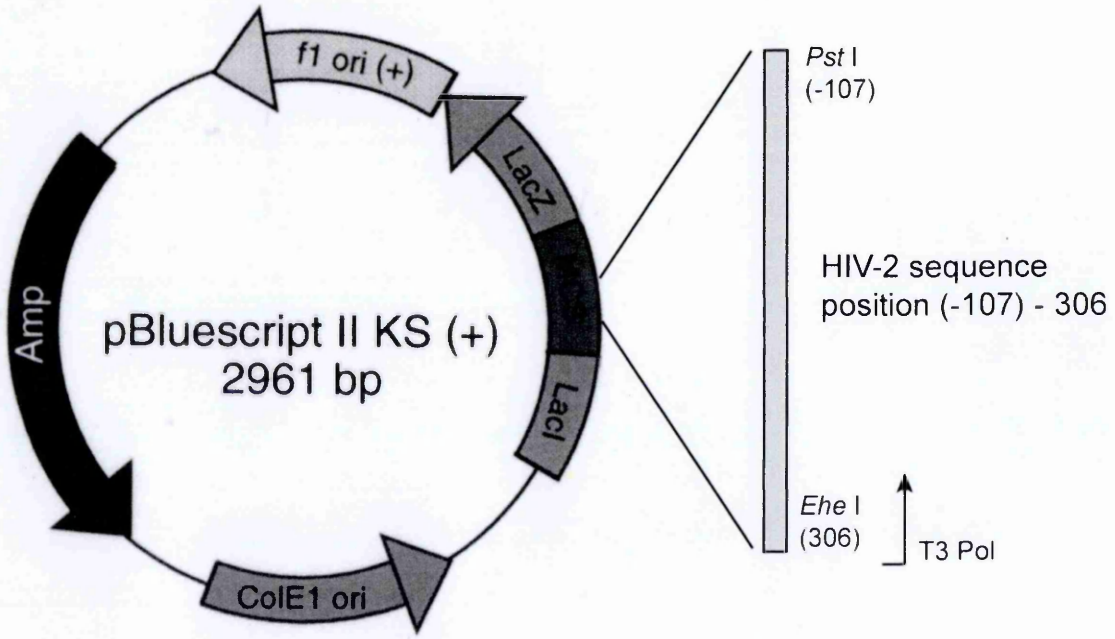

pSVR
12828bp
Pvu II (11973)
Bgl I (11759)
Eco RI (11629)
Hind III (11823)
Nco I (11719)
Eco RI (12073)
Spe I (12518)
Nhe I (12810)
Ehe I (306)
Aat II (11444)
Apa LI (10228)
Apa LI (11244)
Xba I (~9800)
SV40 ori
LTR
gag
EcoRV (1101)
BamHI (1346)
HindIII (1458)
Apa I (1855)
BamHI (1929)
Xho I (2032)
Mst I (2360)
EcoRI (2659)
EcoRV (2939)
AvrII (3014)
Xmn I (3418)
Spe I (3515)
pol
BamHI (3604)
Bam HI (3671)
Bam HI (3708)
Sca I (4585)
Bcl I (4666)
Hinc II (4188)
Xba I (4974)
Xba I (5067)
tat (5845-6140 bp)
rev (6071-6140 bp)
Nsi I (5589)
BstXI (5629)
HindIII (5783)
Ehe I (5864)
Nsi I (6369)
Apa LI (6827)
Fsp I (6585)
Bst XI (6919)
env
RRE (7693-7912 bp)
tat (8307-8430 bp)
rev (8307-8539 bp)
Nco I (7969)
Nco I (8762)
BsmI (8581)
BamHI (8570)
LTR
Nhe I (9479)
Spe I (9187)
Bst XI (2828)
Bst XI (3056)
Bst XI (3086)
Bst XI (4143)


pSVRΔNB
12278 bp
LTR
SV40 ori
gag
pol
env
tat
(8307-8430 bp)
RRE
(7693-7912 bp)
rev
(8307-8539 bp)
tat
(5845-6140 bp)
rev
(6071-6140 bp)
Sal I linker
6919 bp
6369 bp
Pvu II (11973)
Bgl I (11759)
Eco RI (11629)
Hind III (11823)
Nco I (11719)
Eco RI (12073)
Spe I (12518)
Nhe I (12810)
Aat II (11444)
Apa LI (10228)
Apa LI (11244)
Xba I (~9800)
Ehe I (306)
EcoRV (1101)
BamHI (1346)
HindIII (1458)
Apa I (1855)
BamHI (1929)
Xho I (2082)
Mst I (2360)
EcoRI (2659)
EcoRV (2939)
AvrII (3014)
Xmn I (3418)
Spe I (3515)
BamHI (3604)
BamHI (3671)
BamHI (3708)
Sca I (4585)
Hinc II (4188)
Bcl I (4666)
Xba I (4974)
Xba I (5067)
NsiI (5589)
BstXI (5629)
HindIII (5783)
Ehe I (5864)
Nhe I (9479)
Spe I (9187)
Nco I (8762)
BsmI (8581)
BamHI (8570)
Nco I (7969)


ApaI
AatII
11444
EcoRI
11629
HindIII
11823
EcoRI
12073
SV40 ori
f1 ori
SpeI
12518
LTR
NheI
12810
NarI
306
XmnI
ScaI
Amp^r
pGRAXS
6.4kb
gag
EcoRV
1101
HindIII
1458
ApaI
1855
XhoI
2032
NsiI
BstX1
SacI
BamHI
HindIII
ClaI
Csp45
SmaI
KpnI
EcoRI


f1 ori (+)
LacZ
LacI
ColE1 ori
Amp
pBluescript II KS (+)
2961 bp
Ehe I
(306)
HIV-2 sequence
position 306 - 751
Kpn I
(751)
T3 Pol

KS2ΨKE


f1 ori (+)
LacZ
LacI
ColE1 ori
Amp
pBluescript II KS (+)
2961 bp
Pst I
(-107)
HIV-2 sequence
position (-107) - 306
Ehe I
(306)
T3 Pol

KS2ΨEP